



Universidad
de Alcalá

ESCUELA DE DOCTORADO
Servicio de Estudios Oficiales de
Posgrado

DILIGENCIA DE DEPÓSITO DE TESIS.

Comprobado que el expediente académico de D./D^a _____
reúne los requisitos exigidos para la presentación de la Tesis, de acuerdo a la normativa vigente, y habiendo
presentado la misma en formato: ☐ soporte electrónico ☐ impreso en papel, para el depósito de la
misma, en el Servicio de Estudios Oficiales de Posgrado, con el nº de páginas: _____ se procede, con
fecha de hoy a registrar el depósito de la tesis.

Alcalá de Henares a _____ de _____ de 20____



Fdo. El Funcionario



ACTA DE EVALUACIÓN DE LA TESIS DOCTORAL

Año académico 2016/17

DOCTORANDO: **ROMERO GALLARDO, CRISTINA**
DN.I./PASAPORTE: ****363

PROGRAMA DE DOCTORADO: **D432 ECONOMÍA Y GESTIÓN EMPRESARIAL**
DEPARTAMENTO DE: **ECONOMÍA**
TITULACIÓN DE DOCTOR EN: **DOCTOR/A POR LA UNIVERSIDAD DE ALCALÁ**

En el día de hoy 20/09/17, reunido el tribunal de evaluación nombrado por la Comisión de Estudios Oficiales de Posgrado y Doctorado de la Universidad y constituido por los miembros que suscriben la presente Acta, el aspirante defendió su Tesis Doctoral, elaborada bajo la dirección de FRANCISCO J. CALLEALTA BARROSO // .

Sobre el siguiente tema: *FUENTES ESTADÍSTICAS SOBRE LA DISTRIBUCIÓN PERSONAL DE LA RENTA EN ESPAÑA: INDICIOS DE INCONGRUENCIA Y ESTIMACIÓN COMBINADA MEDIANTE PROCESOS DE "MATCHING"*

Finalizada la defensa y discusión de la tesis, el tribunal acordó otorgar la CALIFICACIÓN GLOBAL¹ de (no apto, aprobado, notable y sobresaliente): sobresaliente

Alcalá de Henares, 20 de SEPTIEMBRE de 2017

EL PRESIDENTE


Fdo.: J. Miguel Cerezo

EL SECRETARIO


Fdo.: Mercedes Prieto

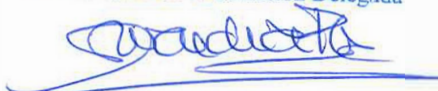
EL VOCAL


Fdo.: Pedro Rouille

Con fecha 4 de octubre de 2017, la Comisión Delegada de la Comisión de Estudios Oficiales de Posgrado, a la vista de los votos emitidos de manera anónima por el tribunal que ha juzgado la tesis, resuelve:

- ☒ Conceder la Mención de "Cum Laude"
☐ No conceder la Mención de "Cum Laude"

La Secretaria de la Comisión Delegada



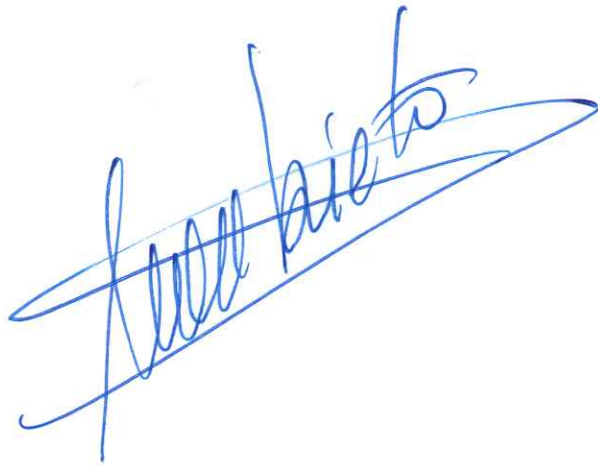
FIRMA DEL ALUMNO,


Fdo.: Cristina Romero

¹ La calificación podrá ser "no apto" "aprobado" "notable" y "sobresaliente". El tribunal podrá otorgar la mención de "cum laude" si la calificación global es de sobresaliente y se emite en tal sentido el voto secreto positivo por unanimidad.

INCIDENCIAS / OBSERVACIONES:

El miembro del Tribunal Administrativo de
Medio Ambiente que actúa
como Vocal Titular ha sido
sustituido por el vocal suplente
Pedro Herrilla Novella, por motivo
vacaciones que le han su pedido
asistir



Fdo: M^c Mercedes Nieto
267



Universidad
de Alcalá

COMISIÓN DE ESTUDIOS OFICIALES
DE POSGRADO Y DOCTORADO

En aplicación del art. 14.7 del RD. 99/2011 y el art. 14 del Reglamento de Elaboración, Autorización y Defensa de la Tesis Doctoral, la Comisión Delegada de la Comisión de Estudios Oficiales de Posgrado y Doctorado, en sesión pública de fecha 4 de octubre, procedió al escrutinio de los votos emitidos por los miembros del tribunal de la tesis defendida por *ROMERO GALLARDO, CRISTINA*, el día 20 de septiembre de 2017, titulada *FUENTES ESTADÍSTICAS SOBRE LA DISTRIBUCIÓN PERSONAL DE LA RENTA EN ESPAÑA: INDICIOS DE INCONGRUENCIA Y ESTIMACIÓN COMBINADA MEDIANTE PROCESOS DE "MATCHING"*, para determinar, si a la misma, se le concede la mención "cum laude", arrojando como resultado el voto favorable de todos los miembros del tribunal.

Por lo tanto, la Comisión de Estudios Oficiales de Posgrado **resuelve otorgar** a dicha tesis la

MENCIÓN "CUM LAUDE"

Alcalá de Henares, 10 de octubre de 2017
EL PRESIDENTE DE LA COMISIÓN DE ESTUDIOS
OFICIALES DE POSGRADO Y DOCTORADO



Juan Ramón Velasco Pérez

Copia por e-mail a:

Doctorando: ROMERO GALLARDO, CRISTINA

Secretario del Tribunal: MARÍA DE LAS MERCEDES PRIETO ALAIZ.

Director de Tesis: FRANCISCO J. CALLEALTA BARROSO



UNIVERSIDAD DE ALCALÁ

FACULTAD DE CIENCIAS ECONÓMICAS, EMPRESARIALES Y TURISMO

DEPARTAMENTO DE ECONOMÍA

Doctorado en Economía Aplicada

**Fuentes Estadísticas sobre la Distribución Personal
de la Renta en España: Indicios de Incongruencia
y Estimación Combinada mediante
Procesos de “*Matching*”**

TESIS DOCTORAL

Cristina Romero Gallardo

Director

Dr. D. Francisco Javier Callealta Barroso

Alcalá de Henares, 2017

*A mis queridos padres,
Alfonso y Pura,
por su esfuerzo, dedicación
y apoyo incondicional*

AGRADECIMIENTOS

Deseo expresar mi eterno y profundo agradecimiento al profesor Dr. D. Francisco Javier Callealta Barroso por haberme brindado la oportunidad de recurrir a su capacidad y conocimiento científico, así como por su labor de dirección, orientación, dedicación e inestimables aportaciones, sin las cuales este trabajo no hubiera llegado a término.

Asimismo, manifiesto mi gran agradecimiento a María del Carmen Ortega Poza, por compartir conmigo sus conocimientos en Contabilidad Nacional.

Mi más sincero agradecimiento también a todos los miembros del Departamento de Economía de esta Universidad, por todo el apoyo recibido.

Finalmente, debo reconocer una inmensa deuda de gratitud con mi querido Pierre, familia, amigos, jefes y compañeros de trabajo, y todos aquellos que, de una forma u otra, han contribuido a que andar este camino fuera posible.

Índice general

Índice de cuadros y figuras	V
Índice de cuadros.....	V
Índice de figuras	VIII

INTRODUCCIÓN GENERAL.....	9
---------------------------	---

CAPÍTULO I: FUENTES ESTADÍSTICAS SOBRE INGRESOS DE LOS HOGARES.....17

1. INTRODUCCIÓN.....	19
2. SOBRE LAS FUENTES ESTADÍSTICAS A NIVEL AGREGADO.....	20
2.1. Rasgos básicos del Sistema Europeo de Cuentas Nacionales y Regionales de la Unión Europea (SEC 2010)	20
2.1.1. Las unidades estadísticas y sus agrupaciones: los sectores institucionales.....	21
2.1.2. Las operaciones económicas y el sistema global de cuentas como marco del proceso de distribución de la renta.....	26
2.1.3. Las Cuentas Regionales.....	32
2.1.3.1. Conceptos básicos del sistema de cuentas regionales	32
2.1.3.2. Las cuentas regionales de los hogares	35
3. SOBRE LAS FUENTES ESTADÍSTICAS A NIVEL INDIVIDUAL	36
3.1. La Encuesta de Condiciones de Vida	37
3.2. La Encuesta de Presupuestos Familiares	41
3.3. La Encuesta Financiera de las Familias.....	44
3.4. Fuentes fiscales.....	47
3.4.1. La Muestra del Impuesto sobre la Renta de las Personas Físicas	48
3.4.2. El Panel de declarantes de IRPF.....	49
4. DISPONIBILIDAD DE LOS DATOS	55

CAPÍTULO II: LOS INGRESOS DE LOS HOGARES: CONGRUENCIA ENTRE LA ENCUESTA DE CONDICIONES DE VIDA Y LA CONTABILIDAD NACIONAL DE ESPAÑA 59 |

1. INTRODUCCIÓN.....	61
----------------------	----

2. LA DISTRIBUCIÓN DE LOS TAMAÑOS DE LOS INGRESOS EN ESPAÑA SEGÚN LA ECV.....	62
3. MARCO TEÓRICO-CONCEPTUAL PARA LA COMPARACIÓN DE INGRESOS DE LOS HOGARES... ..	69
3.1. Análisis de la homogeneidad entre la ECV y la CNE.....	69
3.1.1. Análisis de la homogeneidad entre la ECV y CNE relativa a las variables de ingresos: Cuenta de asignación de la renta primaria.	70
3.1.2. Análisis de la homogeneidad entre la ECV y CNE relativa a las variables de ingresos: Cuenta de distribución secundaria de la renta	92
4. NIVELES DE CONGRUENCIA ENTRE LA ECV Y CNE	114

CAPÍTULO III: EL CONSUMO Y LA SITUACIÓN FINANCIERA DE LOS HOGARES DE LAS ECV

1. INTRODUCCIÓN.....	119
2. ESTUDIO DE RELACIONES TEÓRICAS PARA EL ANÁLISIS DE LA OCULTACIÓN	120
2.1. Determinantes del consumo y el ahorro	124
2.1.1. Factores objetivos.....	126
2.1.2. Factores estructurales	131
3. METODOLOGÍA PROPUESTA PARA ESTIMAR EL CONSUMO Y LA SITUACIÓN FINANCIERA DE LOS HOGARES DE LA ECV	138
3.1. La Fusión Estadística de la BDH (ECV), la EPF y la EFF	139
3.1.1. Fundamentación teórica	139
3.1.2. Fases de la Fusión Estadística de la BDH (ECV), la EPF y la EFF.....	142
3.1.2.1. Definición del papel de cada muestra e identificación de variables	142
3.1.2.2. Armonización y conciliación de fuentes	143
3.1.2.3. Selección de variables de fusión	149
3.1.2.4. Elección del tipo de fusión y del criterio de función de distancia a aplicar.....	154
3.1.2.5. Aplicación y evaluación de resultados	163
4. ESTIMACIÓN DE PARTIDAS NO DISPONIBLES EN LAS FUENTES DE ESTUDIO	166
4.1. Estimación de partidas no disponibles en la CNE Base 2010.....	166
4.2. Estimación de partidas no disponibles en la ECV.....	167
5. EVIDENCIAS DIRECTAS Y CIRCUNSTANCIALES ACERCA DEL FENÓMENO DE LA SUBDECLARACIÓN: INDICADORES DE INCONGRUENCIA.....	175

CAPÍTULO IV: LA ESTIMACIÓN COMBINADA DE LA DISTRIBUCIÓN PERSONAL DE LA RENTA (DPR)	185
1. LA ESTIMACIÓN DE UNA DPR CONGRUENTE: ANTECEDENTES Y OBJETIVOS	187
1.1. El modelo pionero de Pena et al.....	189
1.2. El modelo de Ortega y Callealta como precedente de la utilización de indicadores de rareza para el estudio de la ocultación.....	192
1.3. El modelo de Callealta y Romero, como precedente de la utilización de fusiones estadísticas para el estudio de la ocultación	196
2. MODELO PROPUESTO PARA LA ESTIMACIÓN COMBINADA DE LA DPR.....	198
2.1. Formulación inicial del modelo para la corrección de los microdatos...	198
2.2. Adaptación del modelo para que no dependa de la renta declarada	200
2.3. Metodología propuesta para la detección de observaciones incongruentes y la medida de su grado de incongruencia o propensión a la ocultación	202
2.4. Estimación de mínimos cuadrados de la propensión a la ocultación indeterminada, k.....	207
3. LAS ECV (BASE 2004) Y SU CORRECCIÓN CON EL USO DE REGISTROS ADMINISTRATIVOS EN LAS ECV (BASE 2013): UNA VÍA PARA EL DE ANÁLISIS DE LA OCULTACIÓN DE LOS HOGARES.....	210
3.1. Identificación de los hogares de las ECV en las bases 2004 y 2013 mediante procedimientos de fusión estadística.....	212
3.2. Análisis de las ocultaciones desveladas por las ECV	214
3.2.1. Algunas conclusiones	229
4. IMPLEMENTACIÓN PRÁCTICA DEL MODELO PROPUESTO PARA LA ESTIMACIÓN COMBINADA DE LA DPR.....	231
4.1. Determinación de los niveles de incongruencia de las rentas declaradas de los hogares	232
4.1.1. Indicadores de incongruencia específicos de cada partida	232
4.1.2. Indicadores de incongruencia general.....	238
5. ESTIMACIÓN COMBINADA DE LA DISTRIBUCIÓN PERSONAL DE LA RENTA (DPR) Y ANÁLISIS DE SUS RESULTADOS	239
5.1. Proceso de la Estimación combinada de la DPR (base 2004).....	239
5.1.1. La progresividad estimada “k” de las tasas de ocultación	242
5.2. Proceso de la Estimación combinada de la DPR (base 2013).....	244
5.3. Las estimaciones obtenidas.....	245
5.3.1. Distribuciones personales la renta para el año 2008	245

5.3.2. Distribuciones personales la renta para el año 2011	250
5.4. Resultados comparativos con las ECV originales	253
CONCLUSIONES Y LÍNEAS FUTURAS DE INVESTIGACIÓN	271
BIBLIOGRAFÍA	279
ANEXOS	299
ANEXO 1. PROGRAMAS INFORMÁTICOS PARA LA CONSTRUCCIÓN DE VARIABLES HOMOGÉNEAS EN LOS PROCESOS DE FUSIÓN DE LA ECV-EPF Y ECV-EPF-EFF	301
Anexo 1.a. Programa que prepara los Datos de Personas de las EPF para matching	301
Anexo 1.b. Programa que prepara los Datos de Hogares de las EPF para matching	308
Anexo 1.c. Programa que prepara los Datos de Hogares de las EFF para matching	333
ANEXO 2. PROGRAMA INFORMÁTICO QUE REALIZA EL MATCHING ENTRE LOS ARCHIVOS QUE SE FUSIONAN	368
ANEXO 3. PROGRAMA DE CÁLCULO DE LOS INDICIOS DE INCONGRUENCIA.....	374
ANEXO 4. PROGRAMA QUE CALCULA “K” PARA UNA PARTIDA Y LA CORRIGE.....	377
ANEXO 5. PROGRAMA DE CORRECCIÓN DE LA BDH	380

Índice de cuadros y figuras

Índice de cuadros

Capítulo I: Fuentes estadísticas sobre ingresos de los hogares

Cuadro 1.1 Sectores institucionales en el SEC 2010.....	25
Cuadro 1.2 Presentación sinóptica de las cuentas de los sectores institucionales, saldos contables y principales agregados en el SEC 2010.....	28
Cuadro 1.3 Cuentas de distribución de la renta para el sector hogares en el SEC 2010.....	31
Cuadro 1.4 Cuentas regionales de renta de los hogares en el SEC 2010.....	36
Cuadro 1.5 Disponibilidad de fuentes estadísticas desde 2004.....	57

Capítulo II: Los ingresos de los hogares: Congruencia entre la Encuesta de Condiciones de Vida y la Contabilidad Nacional de España

Cuadro 2.1 Renta disponible neta años 2008 y 2011: CNE vs. ECV Base 2004...63	63
Cuadro 2.2 Renta disponible neta años 2008 y 2011: CNE vs. ECV Base 2013...63	63
Cuadro 2.3 Componentes de la renta disponible del hogar en la ECV.....65	65
Cuadro 2.4 Errores relativos de los componentes de ingreso de la ECV (base 2004) sobre CNE.....66	66
Cuadro 2.5 Errores relativos de los componentes de ingreso de la ECV (base 2013) sobre CNE.....66	66
Cuadro 2.6 Cuentas de renta del sector hogares en el SEC 2010 compatibles con los ingresos de la ECV.....71	71
Cuadro 2.7 Cuenta de asignación de la renta primaria del sector hogares: correspondencia entre la CNE y la ECV.....91	91
Cuadro 2.8 Cuenta de distribución secundaria de la renta del sector hogares: correspondencia entre la CNE y la ECV.....113	113
Cuadro 2.9 Cuenta de la renta de los hogares: ECV vs. CNE.....115	115

Capítulo III: El consumo y la situación financiera de los hogares de las ECV

Cuadro 3.1 El gasto en consumo final de los hogares desde una perspectiva geográfica.....121	121
Cuadro 3.2 El gasto en consumo final y el consumo final efectivo.....122	122
Cuadro 3.3 Estructura de la fusión estadística.....141	141
Cuadro 3.4 Variables homogéneas en la BDH base 2013 y en la EPF.....146	146
Cuadro 3.5 Variables homogéneas en la BDH base 2013-EPF y en la EFF.....147	147
Cuadro 3.5 (cont.) Variables homogéneas en la BDH base 2013-EPF y en la EFF.....148	148

Cuadro 3.6 Selección de variables de matching en hogares de dos adultos sin niños dependientes económicamente para el año de referencia de los ingresos 2008.....	152
Cuadro 3.7 Selección de variables de matching en hogares de dos adultos sin niños dependientes económicamente para el año de referencia de los ingresos 2011.....	153
Cuadro 3.8 Variables matching para el proceso de fusión estadística BDH base 2013 año 2009 vs. EPF año 2008.....	159
Cuadro 3.9 Variables matching para el proceso de fusión estadística BDH base 2013 año 2012 vs. EPF año 2011.....	160
Cuadro 3.10 Variables matching para el proceso de fusión estadística BDH base 2013 año 2009 - EPF año 2008 vs. EF2008.....	161
Cuadro 3.11 Variables matching para el proceso de fusión estadística BDH base 2013 año 2012 - EPF año 2011 vs. EF2011.....	162
Cuadro 3.12 Estadísticos descriptivos del Gasto de los Hogares observado (EPF) e imputado (BDH b2013)	164
Cuadro 3.13 Estadísticos descriptivos de la Riqueza Neta de los Hogares observada (EFF) e imputada (BDH b2013)	165

Capítulo IV: La estimación combinada de la distribución personal de la renta (DPR)

Cuadro 4.1.a Variables comunes de control para la fusión de las bases de las ECV (personas)	213
Cuadro 4.1.b Variables comunes de control para la fusión de las bases de las ECV (hogares)	213
Cuadro 4.2 Tipología de los casos afectados por el fenómeno de la ocultación en un determinado tipo de renta.....	214
Cuadro 4.3 Efecto sobre el Excedente de explotación Neto	217
Cuadro 4.4 Efecto sobre la Renta mixta neta	218
Cuadro 4.5 Efecto sobre los Sueldos y Salarios.....	219
Cuadro 4.6 Efecto sobre las Cotizaciones sociales a cargo del empleador....	220
Cuadro 4.7 Efecto sobre los Intereses, Dividendos y Ganancias de Inversiones de Capital recibidos netos.....	221
Cuadro 4.8 Efecto sobre el Saldo de rentas primarias netas.....	222
Cuadro 4.9 Efecto sobre Impuestos corrientes + Cotizaciones sociales (empleos-recursos)	224
Cuadro 4.10 Efecto sobre Prestaciones Sociales distintas de las transferencias sociales en especie (Recursos)	225
Cuadro 4.11 Efecto sobre Otras transferencias corrientes (Recursos)	226
Cuadro 4.12 Efecto sobre Otras transferencias corrientes (Empleos)	227

Cuadro 4.13 Efecto sobre Renta disponible neta.....	229
Cuadro 4.14.a Indicadores específicos (Cuenta de Asignación de la Renta Primaria)	234
Cuadro 4.14.b Indicadores específicos asociados a las partidas de la Cuenta de Asignación de la Renta Primaria, calculados para la base 2004.....	235
Cuadro 4.14.c Indicadores específicos asociados a las partidas de la Cuenta de Asignación de la Renta Primaria, calculados para la base 2013.....	236
Cuadro 4.15.a Indicadores específicos (Cuenta de Distribución Secundaria de la Renta)	237
Cuadro 4.15.b Indicadores específicos asociados a las partidas de la Cuenta de Redistribución de Rentas, calculados para la base 2004.....	237
Cuadro 4.15.c Indicadores específicos asociados a las partidas de la Cuenta de Redistribución de Rentas, calculados para la base 2013.....	238
Cuadro 4.16.a Indicadores generales.....	238
Cuadro 4.16.b Indicadores generales, calculados para la base 2004.....	239
Cuadro 4.16.c Indicadores generales, calculados para la base 2013.....	239
Cuadro 4.17.a Variación de la pendiente de la tasa de ocultación cuando aumenta G_000.....	240
Cuadro 4.17.b Elección del indicador para estimar la pendiente de la tasa de ocultación.....	241
Cuadro 4.18.a Especificaciones de los modelos de corrección de las partidas de la Cuenta de Asignación de la Renta Primaria, para la base 2004.....	241
Cuadro 4.18.b Especificaciones de los modelos de corrección de las partidas de la Cuenta de Distribución Secundaria de la Renta, para la base 2004.....	242
Cuadro 4.19 Valores estimados para la progresividad "k"	243
Cuadro 4.20.a Especificaciones de los modelos de corrección de las partidas de la Cuenta de Asignación de la Renta Primaria, para la base 2013.....	244
Cuadro 4.20.b Especificaciones de los modelos de corrección de las partidas de la Cuenta de Distribución Secundaria de la Renta, para la base 2013.....	245
Cuadro 4.21.a Congruencia de la Distribución Personal de la Renta Corregida con CNE (2008).....	246
Cuadro 4.21.b Distribución de la Renta Disponible Neta per cápita (2008)	247
Cuadro 4.21.c Distribución de la Renta Disponible Neta per cápita equivalente (2008)	249
Cuadro 4.22.a Congruencia de la Distribución Personal de la Renta Corregida con CNE (2011)	250
Cuadro 4.22.b Distribución de la Renta Disponible Neta per cápita (2011)	251

Cuadro 4.22.c Distribución de la Renta Disponible Neta per cápita equivalente (2011)	252
Cuadro 4.23.a Renta Disponible Neta per cápita equivalente, según Tipo de Hábitat del Hogar (2008)	255
Cuadro 4.23.b Renta Disponible Neta per cápita equivalente, según Tipo de Hábitat del Hogar (2011).....	256
Cuadro 4.24.a Renta Disponible Neta per cápita equivalente, según Categoría socio-económica del Sustentador Principal del Hogar (2008)	257
Cuadro 4.24.b Renta Disponible Neta per cápita equivalente, según Categoría socio-económica del Sustentador Principal del Hogar (2011)	258
Cuadro 4.25.a Renta Disponible Neta per cápita equivalente, según Grupo de Edad del Sustentador Principal del Hogar (2008)	259
Cuadro 4.25.b Renta Disponible Neta per cápita equivalente, según Grupo de Edad del Sustentador Principal del Hogar (2011)	260
Cuadro 4.26.a Renta Disponible Neta per cápita equivalente, según Género del Sustentador Principal del Hogar (2008)	261
Cuadro 4.26.b Renta Disponible Neta per cápita equivalente, según Género del Sustentador Principal del Hogar (2011)	262
Cuadro 4.27.a Renta Disponible Neta per cápita equivalente, según Nivel de Estudios del Sustentador Principal del Hogar (2008)	264
Cuadro 4.27.b Renta Disponible Neta per cápita equivalente, según Nivel de Estudios del Sustentador Principal del Hogar (2011)	265
Cuadro 4.28.a Renta Disponible Neta per cápita equivalente, según Nacionalidad del Sustentador Principal del Hogar (2008)	267
Cuadro 4.28.b Renta Disponible Neta per cápita equivalente, según Nacionalidad del Sustentador Principal del Hogar (2011)	268

Índice de figuras

Capítulo I: Fuentes estadísticas sobre ingresos de los hogares

Figura 1.1 La renta del sector hogares y sus componentes en el SEC 2010.....	30
--	----

Capítulo IV: La estimación combinada de la distribución personal de la renta (DPR)

Figura 4.1 Clases inducidas por una relación ordinal para cada caso (xi, zi)	203
Figura 4.2 Identificación de clases de interés para la construcción de indicadores de incongruencia	205

Introducción general

El contexto general de esta investigación se enmarca en el ámbito de estudio de la distribución de la renta. Por un lado, se tratará de analizar la congruencia y complementariedad de las fuentes estadísticas que proporcionan información sobre los ingresos de los hogares y sus miembros y, por otro, se profundizará en la utilización de éstas con el objetivo de llevar a cabo la estimación combinada de una distribución del tamaño de los ingresos¹ para el caso específico de los hogares españoles, tratando de armonizar las principales discrepancias que en dichas fuentes se observen.

No podemos abordar el estudio de la distribución de la renta sin antes esbozar algunas nociones sobre este concepto. En general, se entiende como *distribución de la renta* la manera en la que se reparte el conjunto de recursos materiales derivados de la actividad económica de acuerdo con distintos parámetros o perspectivas de referencia. La distribución de la renta puede analizarse desde cuatro perspectivas: sectorial, espacial, funcional y personal.

Desde una perspectiva *sectorial*, la distribución de la renta se refiere al reparto de la misma entre los diferentes sectores productivos (agricultura, industria, construcción y servicios), que componen el Producto Interior Bruto como indicador de actividad económica. El análisis desde esta óptica estaría indicado, por ejemplo, para el diseño de políticas de oferta.

Desde una perspectiva *espacial*, alude a las diferencias que pueden existir en los niveles de renta desde un punto de vista territorial y, en consecuencia, es objeto de análisis particularizado por parte de la economía regional.

Desde un enfoque *funcional*, la distribución de la renta se refiere a la forma en que ésta se reparte entre todos y cada uno de los factores de producción. Este enfoque ya fue contemplado por los autores clásicos, quienes distinguían entre terratenientes, capitalistas y trabajadores. David Ricardo estableció una diferencia clara entre las rentas de la tierra, los beneficios y los salarios, apuntando que la renta de la tierra y los beneficios se obtenían como residuo, y que la distribución del producto dependía fundamentalmente de la cantidad y calidad de las tierras de cultivo, la tecnología aplicada y el fondo salarial y el número de trabajadores. Por su parte, los autores post-ricardianos introdujeron el concepto de contribución marginal de cada factor de producción al producto final; así, la aportación de cada factor de producción estaba ligada a su productividad marginal. Otras aportaciones posteriores han destacado los cambios en la cantidad de factores, su elasticidad de sustitución y el papel de los cambios tecnológicos a la hora de explicar esa distribución funcional.

¹ La expresión "distribución del tamaño de los ingresos", traducción literal de las expresiones inglesas "income size distribution" o "size distribution of income", ampliamente utilizada en la literatura científica para referir a la distribución estadística (empírica o probabilística) de los microdatos de ingresos, no parece arraigar en el ámbito económico-científico español. En su lugar, se emplea con mucha frecuencia la expresión "distribución personal de la renta", incluso cuando las unidades estadísticas realmente observadas no sean personas sino hogares, ya que la expresión más simple "distribución de la renta" generalmente es más empleada en relación con los agregados macroeconómicos de las Cuentas Nacionales. Así, la expresión "distribución de la renta de los hogares" suele hacer pensar más en los agregados de la Cuenta de Rentas del Sector de Hogares de la Contabilidad Nacional, que en la distribución empírica o probabilística del tamaño de las rentas que reciben individualmente los hogares como unidades estadísticas. Es por ello por lo que, a lo largo de esta Memoria, hemos empleado preferentemente la expresión "distribución personal de la renta" para referirnos a cualquier "distribución del tamaño de los ingresos", sean éstos realmente personales o de los hogares en los que se integran.

Finalmente, la distribución *personal* de la renta, desde la que abordaremos este trabajo, se refiere al reparto de la misma entre los individuos o las familias de la sociedad, una vez descontados los impuestos y cotizaciones sociales y añadidas las transferencias económicas que se realicen a los hogares por los poderes públicos (prestaciones sociales y otras transferencias corrientes). Para aislar esta perspectiva de las otras tres, debe tenerse en cuenta que su estudio es independiente del sector en que se produzcan (enfoque sectorial), de la región en que se obtengan (espacial), o de los factores que remunere (enfoque funcional).

En la actualidad, muchos autores que se dedican al estudio de la distribución de la renta se centran en este último enfoque, profundizando no sólo en la descripción y análisis de las diferencias personales de la renta observadas empíricamente, sino incluso, más aún, en su modelización teórica. Así, en las últimas décadas han aparecido un gran número de trabajos que, empezando en Becker (1964) y siguiendo en Mincer (1970), Chiswick (1974), Tinbergen (1975), Taubman (1975), Atkinson (1975), Phelps-Brown (1977), Ward (1978), Howard (1979), Wright (1979) y Lydall (1979), entre otros, han intentado desarrollar y verificar empíricamente una teoría general de la distribución de la renta entre individuos. Más recientemente, en España, autores como Garde *et al.* (1996), Ayala *et al.* (1996), Cantó *et al.* (2000), etc. miden y analizan las diferencias entre clases sociales; estudian los efectos de las políticas redistributivas; aportan propuestas sobre el diseño de políticas públicas dirigidas a reducir la desigualdad, la polarización en la distribución de la renta, la pobreza o la exclusión social; analizan la importancia de la distribución de la renta en el crecimiento económico; y revisan las nuevas aportaciones metodológicas que surgen continuamente.

Sin embargo, las dificultades presentes en la recogida de información hacen que a menudo se aprecien inconsistencias importantes entre las diferentes fuentes. Concretamente, trabajos como los realizados por Alcaide Inchausti, A. y J. (1974), (1977) y (1983), Pena *et al.* (1996), Prieto y Pena (2000), Andrés y Mercader (2001), Pascual y Sarabia (2006), etc. han venido evidenciando a lo largo del tiempo estas discrepancias entre ciertas fuentes de información concretas (Encuestas Básicas de Presupuestos Familiares, Panel de Hogares de la Unión Europea, Contabilidad Nacional...), y para distintos períodos (1964, 1967, 1970, 1973, 1981, 1991, 2000...).

En el panorama internacional, aportaciones como las de Moore *et al.* (2000), Nordberg *et al.* (2001), Ravallion (2001), Meyer *et al.* (2008), ponen también de manifiesto ciertas divergencias presentes en las variables de índole económico pertenecientes a los hogares cuando se ponen en comparación las distintas fuentes dedicadas al estudio de aquéllas.

Las políticas públicas deben encaminarse a reducir las desigualdades económicas entre los individuos y a hacer la sociedad un poco más justa cada día. Y para poder realizar una asignación adecuada de los recursos públicos en dicha dirección, es preciso conocer la situación real de bienestar de los individuos, lo que depende de forma importante de los ingresos que realmente perciben y de su nivel de riqueza. Por tanto, aquellos defectos en la información que suministran las fuentes estadísticas con base a las que se planifican tales políticas pueden producir errores a la hora de asignar aquellos recursos, como los que se deberían asignar a los más desfavorecidos en forma

de prestaciones sociales tales como ayudas para estudios, pensiones de jubilación, etc. alejando los efectos de tales políticas de sus objetivos deseados, o haciéndolas menos eficientes de lo que podrían ser, cuando menos. En cualquier caso, la constatación de la existencia de diferencias considerables en las fuentes estadísticas califica negativamente la deseada calidad de la información, la cual es esencial para que el análisis de las clases y que los efectos de las políticas sobre ellas sean los correctos.

Así pues, constatada, por una parte, la sistemática existencia de discrepancias considerables entre la renta disponible de los hogares y sus diferentes partidas o tipos de rentas que se desprenden de las principales encuestas oficiales y de las que publica la Contabilidad Nacional de España (CNE); y, por otra, la gran heterogeneidad en los niveles de las discrepancias presente entre los agregados de las diversas partidas de rentas de los hogares obtenidos con las encuestas oficiales y los de la CNE; Callealta y Romero (2013a) advierten también de que, nutriéndose las rentas percibidas en los diferentes grupos socio-económicos presentes en la sociedad, fundamentalmente, de distintos tipos de rentas (salariales, por cuenta propia, jubilaciones, desempleo,...), dicha heterogeneidad en los niveles de discrepancia de las diferentes partidas de rentas podría afectar, muy posiblemente también de forma heterogénea, a la estimación de las rentas en dichos grupos, influyendo esto consecuentemente en la apreciación de algunas de sus características de máximo interés. Así, por ejemplo, se observan tasas de pobreza extrañamente altas en algunos grupos socioeconómicos cuyas rentas fundamentales se nutren precisamente de las partidas de rentas que mayor discrepancia presentan (por infradeclaración) con las correspondientes que muestra la CNE; lo que les hace pensar si realmente todos los pobres de estos colectivos son realmente pobres, o simplemente están clasificados como pobres porque ocultan buena parte de sus rentas.

Con base en lo expuesto anteriormente, el principal objetivo de esta tesis será desarrollar una metodología que, partiendo de los microdatos sobre ingresos que se manifiesten más completos y fidedignos, y considerando como marco de referencia a respetar los agregados nacionales de las rentas más extensamente admitidos, permita estimar una distribución personal de la renta que consensúe de alguna forma objetiva la información discrepante que dichas fuentes de microdatos pudieran presentar. Para ello, echaremos mano también de otras fuentes estadísticas que aporten otros matices sobre los ingresos y/u otros aspectos económicamente relacionados con ellos, como podrían ser el consumo, la riqueza o la situación financiera de las unidades en estudio. A la distribución personal de la renta resultante de esta metodología la denominaremos en el contexto de este trabajo "estimación combinada de la distribución personal de la renta".

Para precisar los objetivos parciales en que estructuraremos la presente investigación, conducentes, en definitiva, a mitigar la incertidumbre y la presumible falta de realismo que las discrepancias entre las mencionadas fuentes inducen sobre las estimaciones de la distribución personal de la renta, adoptaremos en la presente investigación las siguientes hipótesis de trabajo:

1. Las fuentes de microdatos, tanto las que proceden directamente de registros administrativos como las que proceden de encuestas, adolecen de determinados errores de observación ajenos al muestreo,

especialmente en las variables de ingresos, que las hacen discrepar de la realidad. En la mayoría de los casos, la natural reticencia humana (y por tanto, de muchos encuestados) a comunicar a los demás sus datos personales de tipo económico, puede conducir a una recogida de información no siempre fiel a la realidad de las personas y de sus hogares.

2. Las causas de discrepancia entre las fuentes actúan de forma heterogénea sobre las unidades estadísticas analizadas, motivando que la calidad de los microdatos sea muy diferente en clases o colectivos distintos, y sesgando los análisis que se realizan sobre éstos sin tener en cuenta esta circunstancia. Es evidente que la naturaleza de los ingresos de los hogares no es idéntica en toda la población, sino que, para diferentes grupos poblacionales, se asocia fuertemente con ciertas características comunes. Así, se precisa ser asalariado en algún momento para percibir rentas salariales del trabajo; o ser empresario o profesional libre para recibir rentas del trabajo por cuenta propia; o estar parado para percibir prestaciones sociales por desempleo; etc. Y, por tanto, niveles de ocultación en las declaraciones de determinados tipos de ingresos que pudieran ser muy diferentes, o desequilibrados, para los diferentes grupos poblacionales, podrían producir sesgos presumiblemente importantes en las mediciones de fenómenos sociales, como la desigualdad o la pobreza, mediante indicadores métricos objetivos. De este modo, y en caso extremo, si en una cierta subpoblación se ocultaran sistemáticamente la totalidad de sus ingresos reales, declarando ingresos nulos, todos los pertenecientes a dicha subpoblación se posicionarían con práctica seguridad bajo los umbrales de pobreza calculados para la población en su conjunto por métodos objetivos. Sin embargo, los integrantes de tal subpoblación no tienen por qué ser necesariamente pobres; simplemente han ocultado sus rentas. Pero no es éste el único efecto que se produciría; la ocultación de rentas de esta subpoblación sesgaría a la baja los indicadores de posición central de la distribución de ingresos (medias, mediana) y, en consecuencia, sesgará también a la baja el umbral de la pobreza, expulsando de la categoría de pobres a otros perceptores de rentas bajas, verdaderos pobres si declararon sus rentas reales.
3. Las Contabilidades Nacional y Regional, que nos ofrecen datos de los ingresos de los hogares en España y a nivel de comunidad autónoma, respectivamente, a nivel agregado, y cuya información se produce vigilando la coherencia de la misma en varias direcciones, incluyendo la contrastación con la información que al respecto muestran las encuestas oficiales, deben producir resultados agregados más acordes con la realidad que las fuentes de microdatos consideradas individualmente. Esto las hace candidatas naturales a ser utilizadas como referencia para el control de los agregados de la distribución personal de la renta mejorada que pretendemos estimar.

Y, con base en las citadas hipótesis, el primer objetivo que nos planteamos abordar en este trabajo y que acometeremos en el capítulo I del mismo, consiste en el análisis pormenorizado de las fuentes estadísticas de ingresos de los hogares que tradicionalmente han permitido su medición y análisis, tanto desde el enfoque macro, que tiene sus raíces en las cuentas nacionales y, en particular, en las normas basadas en la contabilidad y establecidas en el Sistema Europeo de Cuentas (SEC) -para los países que integran la Unión

Europea-, como desde una aproximación micro, que tiene su origen en la Microeconomía y, en concreto, en el interés por el estudio de la distribución personal de la renta y sus características, como la pobreza y la desigualdad, así como los efectos que las políticas económicas tienen sobre dicha distribución y sus características en los diferentes grupos socio-económicos de la población. Este análisis nos permitirá determinar la fuente que proporciona un mayor grado de detalle sobre la renta de los hogares a nivel de microdatos -como veremos, la Encuesta de Condiciones de Vida del INE (en adelante, ECV)- y condicionará el período de estudio que adoptaremos, procurando la coexistencia de diversas fuentes que nos podrán ser de utilidad en este trabajo para alcanzar nuestros objetivos.

Posteriormente, en el capítulo II de esta investigación abordaremos un segundo de los objetivos, que consistirá en evaluar detalladamente la congruencia, o grado de acuerdo existente, entre los datos de renta proporcionados por la ECV, y la información agregada para los hogares que, de forma oficial, suministra la Contabilidad Nacional de España (CNE). Para ello, examinaremos la *renta disponible* proporcionada por la ECV a nivel agregado tratando de encontrar rasgos que nos orienten en el análisis de las discrepancias observadas con la CNE. Posteriormente, y para profundizar en dicho análisis, analizaremos pormenorizadamente las componentes de la Cuenta de Rentas del Sector Hogares (CRSH) que calcula la CNE y las compararemos con las variables que recogen las ECV, en el ánimo de detectar sus compatibilidades y diferencias. Y, con base en este análisis, diseñaremos y construiremos una CRSH compatible para ambas fuentes que, con cierto nivel de desagregación, nos permitirá comparar los datos de la ECV con los de la CNE, procediendo a calcularla para ambas fuentes y estudiar la congruencia o discrepancias que muestren sus resultados.

Más tarde, y con el objetivo de ir caminando hacia la obtención de la estimación combinada de la distribución personal de los ingresos de los hogares como expresión más acorde con el marco de referencia y pretendidamente más realista que la que podríamos derivar directamente de las ECV, habida cuenta de que el nivel de renta de un hogar permite el consumo por parte de sus miembros y contribuye a los cambios en la riqueza o patrimonio neto de los mismos, en el capítulo III de este trabajo indagaremos en las relaciones derivadas de la Teoría económica que vinculan a la renta con estas y otras variables económicas relacionadas y de índole estructural (como la edad, el sexo, el tamaño del hogar, el nivel de educación, etc.). El fin no será otro que determinar razonadamente con qué información de esta índole deberíamos complementar la información que sobre los hogares proporciona las ECV y de qué otras fuentes de microdatos podríamos obtenerla mediante procesos de fusión estadística para poder analizar el grado de cumplimiento o incumplimiento de aquellas relaciones en los datos económicos declarados por los hogares. Encontraremos información complementaria útil al respecto en la Encuesta de Presupuestos Familiares (EPF) y la Encuesta Financiera de las Familias (EFF). Una vez realizada la fusión de estas fuentes, esta gran cantidad de información complementaria adicional nos servirá de base para diseñar un conjunto de indicadores que, basados en dichas relaciones teóricas estudiadas entre las variables de ingresos, gastos, situación financiera y riqueza de los hogares, así como en otras pruebas circunstanciales derivadas de las propias sinergias de la fusión

de la información, nos permitirán detectar indicios de ocultación o subdeclaración de ingresos, que serán básicos para corregir la subdeclaración detectada a nivel de hogar en la ECV con respecto a la que de forma agregada se publica en los datos oficiales de la CNE; lo que abordaremos finalmente en el capítulo IV.

Así, en el último y cuarto capítulo de esta Memoria, con el objetivo final de realizar la estimación a partir de las ECV de una distribución del tamaño de los ingresos congruente con los datos de CNE, la combinada de la distribución, proponemos un procedimiento flexible y sencillo de corrección de sus microdatos, con base en algunos trabajos anteriores de este doctorando y su director de tesis, y en otros de éste junto a otros autores. La principal novedad del procedimiento propuesto consistirá en hacer depender la presumible ocultación estimada de los hogares, no de sus ingresos declarados como venían planteando los modelos precursores, sino de los indicios objetivos de la incongruencia que pudieran observarse en sus declaraciones y que pudieran informarnos sobre su posible propensión a la ocultación. Para encontrar estos posibles indicios de incongruencia en cada tipo de renta declarada se utilizarán las leyes económicas y/o empíricas estudiadas en el capítulo anterior. Y, basándonos en éstas, se diseñará una metodología general para la construcción de indicadores que medirán los grados de cumplimiento por parte de cada uno de los hogares, y que serán empleados para medir los niveles potenciales de incongruencias (ocultación o sobredeclaración) en las cifras declaradas por aquellos en cada tipo de renta. Con estos indicadores, procederemos a plantear convenientemente el modelo finalmente empleado para la estimación combinada de la distribución personal de la renta en los hogares españoles. Quedando en este modelo un cierto grado de subjetividad para la determinación de uno de sus parámetros, hemos indagado algo más en el fenómeno de la subdeclaración, gracias a la oportunidad que nos brinda el cambio de metodología que el INE ha materializado en la ola del año 2013 de las ECV, publicando para varios años simultáneamente la versión original de las ECV, en la base 2004, y la versión corregida posteriormente con la ayuda de registros administrativos, en la base 2013. A partir de éstas, se ha podido obtener información sobre declaraciones y ocultaciones desveladas en cada uno de los hogares, también mediante procedimientos de fusión estadística de ambas bases de esta encuesta; y, con base en ellas, se dará una propuesta para estimar de forma objetiva un valor plausible para dicho parámetro.

Capítulo I

Fuentes estadísticas sobre ingresos de los hogares

1. Introducción

La medición y análisis de los recursos económicos disponibles para la población tiene una larga tradición que ha dado lugar a medidas coherentes, armonizadas y actualizadas periódicamente de los recursos económicos. Dentro de esta tradición, han surgido dos enfoques principales para medir el ingreso de los hogares (UN, 2011)²:

- El enfoque macro, que tiene sus raíces en las cuentas nacionales y, en particular, las normas contables establecidas en el Sistema de Cuentas Nacionales (SCN)³.
- El enfoque micro, que tiene sus raíces en la microeconomía y, particularmente, en el estudio de las características de la distribución personal de la renta, como la desigualdad y la pobreza, así como de su efecto sobre los diferentes grupos socioeconómicos dentro de la sociedad.

Desde el principio, los economistas y los estadísticos han sido conscientes de las limitaciones de usar las medidas de ingreso agregado como variables proxy del bienestar. Marshall y Pigou explicaron que el "bienestar económico" es sólo un componente del bienestar general, o simplemente, bienestar. Simon Kuznets, uno de los pioneros del SCN, en su primer informe al Congreso de los Estados Unidos en 1934, señaló que "el bienestar de una nación, por lo tanto, apenas puede deducirse de una medida de ingreso nacional..." (Kuznets, 1934). Del mismo modo, el SCN reconoce explícitamente las limitaciones del PIB como medida del bienestar: "No se puede esperar que los movimientos del PIB sean buenos indicadores de cambios en el bienestar general a menos que todos los demás factores que influyen en el mismo permanezcan constantes, pero la historia muestra que nunca es así" (UN, 2011, párrafo 1.69).

Los responsables políticos nunca se han centrado exclusivamente en el crecimiento del PIB como la métrica única para medir el bienestar. Tratan de mejorar el bienestar general de los ciudadanos teniendo en cuenta una serie de factores que van más allá del valor total de los bienes y servicios producidos por un país en un año dado, para incluir preocupaciones distributivas y calidad ambiental. Además, los economistas y los estadísticos siguen estudiando formas de mejorar las medidas de bienestar, incluidas las medidas de bienestar económico.

Los datos del SCN, o macro datos, proporcionan agregados sectoriales compilados a partir de muchas fuentes y presentados dentro de un marco amplio que permite trazar las relaciones entre las distintas cuentas y los sectores institucionales de la economía. En general, los datos del SCN sólo proporcionan información agregada para el sector de los hogares en su conjunto, razón por la cual se puede hacer un mayor uso de fuentes de datos parciales, imputación o estimación. Los datos del SCN también tienen la ventaja de ser totalmente coherentes, sin embargo, no suele ser posible

² UNECE (2011), *The Canberra Group Handbook on Household Income Statistics*, Second Edition, ECE/CES/11, Ginebra.

³ En España, el marco de referencia vigente en materia de Contabilidad Nacional es el Sistema Europeo de Cuentas Nacionales y Regionales de la Unión Europea, 2010 (SEC 2010), coherente con las directrices internacionales sobre Contabilidad Nacional establecidas en el Sistema de Cuentas Nacionales de 2008 (SCN 2008).

utilizarlos para examinar la distribución personal de los recursos dentro de un país.

Por su parte, los microconjuntos de datos pueden utilizarse para analizar no sólo los niveles (agregados), sino también la distribución del ingreso, el consumo y la riqueza en la población, en los diferentes subgrupos y a lo largo del tiempo. Los microdatos pueden proporcionar información importante sobre el bienestar económico de la persona típica; por ejemplo, mientras que el ingreso familiar promedio y la riqueza han aumentado sustancialmente en muchos países durante las últimas dos décadas, no todos los hogares en estos países se han visto beneficiados, lo que refleja el aumento de las desigualdades. Por esta razón, tanto la OCDE (2011) como Stiglitz *et al.* (2009) sostuvieron que cuando se evalúa el bienestar, es mejor enfatizar la perspectiva de los hogares y los individuos, en lugar de las condiciones agregadas de la economía, ya que puede haber una discrepancia entre la situación económica de toda la economía y el bienestar material de los hogares.

Con el objetivo de poder efectuar una comparación sobre la medición de la renta de las familias entre ambos enfoques, en este primer capítulo, profundizaremos en las diferentes fuentes estadísticas que contienen dicha información tanto de forma agregada como a nivel microeconómico.

2. Sobre las fuentes estadísticas a nivel agregado

El objetivo de este apartado es el estudio de las fuentes estadísticas de rentas del sector hogares que nos ofrecen información a nivel agregado: las cuentas económicas del sector hogares que, como a continuación veremos, tenemos disponibles tanto para el conjunto de la economía nacional (Contabilidad Nacional de España -CNE-), como desagregadas por región (Contabilidad Regional).

Como ya anticipábamos en la introducción de este trabajo, para realizar la comparación de los datos de ingresos de los hogares que nos proporcionan las encuestas con la información procedente del Sistema de Contabilidad Nacional, es preciso conocer el marco contable de referencia que subyace bajo el mismo, en la actualidad Sistema Europeo de Cuentas Nacionales y Regionales 2010 (SEC 2010). A continuación, pasamos a explicar brevemente, el contenido, formato y las definiciones que le son aplicables.

2.1. Rasgos básicos del Sistema Europeo de Cuentas Nacionales y Regionales de la Unión Europea (SEC 2010)

A grandes rasgos, podemos decir que la Contabilidad Nacional es una presentación cuantitativa de los flujos económicos que se realizan en un país durante un período determinado, en general, un año.

En España y desde el año 1967, el organismo encargado de la elaboración de las cuentas nacionales es el Instituto Nacional de Estadística (INE). Desde entonces, dicho organismo ha empleado diferentes marcos metodológicos para su obtención: en los primeros años, el Sistema de Contabilidad Nacional de Naciones Unidas y, a partir de 1976, utilizando las diversas versiones del Sistema Europeo de Cuentas propuesto por la Oficina de Estadística de la Unión Europea (EUROSTAT). Hasta 1995 se trabajó bajo el marco "Sistema Europeo de Cuentas Económicas Integradas" (SEC 79)⁴, y a partir de entonces, el INE ha elaborado la Contabilidad Nacional de España bajo el cumplimiento de las directrices marcadas por el "Sistema Europeo de Cuentas Nacionales y Regionales 1995" (SEC 95)⁵, debido a su carácter de reglamento comunitario. Desde la entrada en vigor de la más reciente normativa, septiembre de 2014, la Contabilidad Nacional se calcula bajo el cumplimiento del SEC 2010⁶, el cual constituye una revisión del anterior.

Para una adecuada comprensión de los conceptos contables mencionados en el desarrollo de este trabajo nos interesa tener una visión general de la metodología empleada para la elaboración de las cuentas nacionales en el ámbito de la Unión Europea según la normativa en vigor en el período de estudio: el Sistema Europeo de Cuentas Nacionales y Regionales 2010 (SEC 2010).

El sistema SEC 2010, al igual que su predecesor, el SEC 95, está formado por tres tipos principales de elementos⁷:

- Las unidades (y agrupaciones de éstas) que constituyen el sujeto de las actividades económicas.
- Las variables económicas, que según se refieran a un período de tiempo determinado o a un momento concreto, pueden ser, respectivamente, "flujos" o "stocks".
- Las relaciones entre variables y/o unidades económicas, relaciones que se presentan en un "sistema de cuentas".

A continuación, se realiza una breve descripción de estos tres tipos de elementos, enfatizando los aspectos más interesantes para el análisis de la distribución de la renta, tema central de este trabajo de investigación.

2.1.1. Las unidades estadísticas y sus agrupaciones: los sectores institucionales

La economía de un país es un sistema a través del cual las instituciones y las personas interactúan mediante intercambios y transferencias de bienes, servicios y medios de pago (por ejemplo, dinero), para la producción y el consumo de bienes y servicios.

⁴ Comisión (Eurostat): «Sistema Europeo de Cuentas Económicas Integradas (SEC)», segunda edición, Oficina Estadística de las Comunidades Europeas, Luxemburgo, 1979.

⁵ Reglamento (CE) N° 2223/96 del Consejo de 25 de junio de 1996 relativo al Sistema Europeo de Cuentas Nacionales y Regionales de la Comunidad.

⁶ Reglamento (UE) n° 549/2013 del Parlamento Europeo y del Consejo, de 21 de mayo de 2013, relativo al Sistema Europeo de Cuentas Nacionales y Regionales de la Unión Europea.

⁷ Ambos diferencian también un cuarto elemento que es el "marco input/output" (la versión ampliada de las conocidas tablas input/output) pero que puede considerarse como una forma adicional de presentación de los flujos económicos, incluíble en sentido amplio en el tercer tipo de elementos (el conjunto de cuentas).

Definir las unidades que se utilizan en la Contabilidad Nacional cumple diversos objetivos. En primer lugar, las unidades constituyen los componentes esenciales para definir las economías en términos geográficos, como por ejemplo naciones, regiones y conjuntos de naciones como las uniones monetarias o políticas. En segundo lugar, constituyen los componentes esenciales para agrupar las unidades en sectores institucionales. En tercer lugar, son fundamentales para decidir qué flujos y stocks se registran. Las operaciones entre distintas partes de una misma unidad institucional, en principio, no se registran en las cuentas nacionales.

En el SEC 2010 se utilizan diferentes tipos de unidades que corresponden a tres formas de subdividir la economía:

- Para el análisis de los flujos vinculados a la renta (distribución y redistribución), las operaciones financieras y el patrimonio, las cuentas nacionales clasifican a los agentes económicos en las denominadas "unidades institucionales", de acuerdo con su función o tipo de actividad económica principal. Por agregación, se tienen los denominados "sectores institucionales": hogares, sociedades, administraciones públicas...
- Para describir los procesos de producción y llevar a cabo el análisis input-output se toman las unidades que realizan la producción –unidades de actividad económica a nivel local⁸- y que por agregación llevan a las "ramas de actividad" (rama agraria, rama industrial, etc.). Una actividad se caracteriza por unos insumos de bienes y servicios, un proceso de producción y unos productos (bienes y servicios) obtenidos.
- Y para realizar análisis regionales, hacen falta también unidades que reflejen actividades económicas a nivel local.

Las unidades que constituyen la economía de un país, cuyos flujos y stocks se registran en el SEC 2010, son las que son residentes. La unidad institucional tiene su residencia en un país cuando tiene su centro de interés económico predominante en el territorio económico de tal país. Tales unidades se denominan unidades residentes, independientemente de su nacionalidad, de su forma jurídica o de su presencia en el territorio económico correspondiente en el momento en que efectúen una operación.

El territorio económico consiste en:

- a) la zona (territorio geográfico) sometida a la gestión y el control efectivos de una administración única;
- b) otras zonas francas, incluidos los depósitos y fábricas bajo control aduanero;
- c) el espacio aéreo nacional, las aguas territoriales y la plataforma continental situada en aguas internacionales sobre la que el país disfruta de derechos exclusivos;
- d) los enclaves territoriales, que son territorios geográficos situados en el resto del mundo y utilizados, en virtud de tratados internacionales o de acuerdos entre Estados, por las administraciones públicas del país (como embajadas, consulados, bases militares, bases científicas, etc.);

⁸ Cuando las unidades institucionales desarrollan más de una actividad, es preciso dividir las según el tipo de actividad. Las unidades de actividad económica locales permiten realizar esa presentación.

e) los yacimientos (petróleo, gas natural, etc.) situados en aguas internacionales fuera de la plataforma continental del país, explotados por unidades residentes en el territorio, tal como se han definido en las anteriores letras a) a d).

En el SEC 2010, los barcos de pesca, otros buques, plataformas flotantes y aeronaves reciben el mismo tratamiento que los demás equipos móviles propiedad de y/o explotados por unidades residentes, o propiedad de no residentes y explotados por unidades residentes en el país. Las operaciones relativas a la propiedad (formación bruta de capital fijo) y a la utilización (alquiler, seguros, etc.) de equipos móviles se atribuyen a la economía del país del que son residentes, respectivamente, el propietario y/o quien explota el equipo. En el caso del arrendamiento financiero, se considera que se ha producido un cambio de propiedad.

El territorio económico no incluye enclaves extraterritoriales. Tampoco se incluyen las partes del territorio geográfico del país utilizadas por organizaciones externas tales como administraciones públicas de otros países; instituciones y organismos de la Unión Europea; y organizaciones internacionales en virtud de tratados internacionales entre Estados.

Los territorios utilizados por las instituciones y organismos de la Unión Europea y por las organizaciones internacionales constituyen territorios económicos diferenciados. Una característica de estos territorios es que los únicos residentes son las instituciones.

El término «centro de interés económico predominante» indica que, en el territorio económico de un país, existe un lugar en el cual una unidad realiza actividades y operaciones económicas de cierta amplitud, bien de forma indefinida, bien durante un período determinado, pero suficientemente largo (un año o más). El mero hecho de poseer terrenos o edificios en el territorio económico se considera suficiente para que el propietario tenga un centro de interés económico predominante en dicho territorio.

Las empresas suelen estar vinculadas a una sola economía. La fiscalidad y demás requisitos jurídicos dan lugar generalmente a la utilización de una persona jurídica diferenciada para las operaciones realizadas en cada jurisdicción. Además, se identifica una unidad institucional diferenciada con fines estadísticos cuando una única persona jurídica ejerce actividades sustanciales en dos o más territorios (por ejemplo, en el caso de sucursales, propiedad de terrenos y empresas multiterritoriales). A causa de la división de esas personas jurídicas, está claro el lugar de residencia de cada una de las nuevas empresas identificadas. El término «centro de interés económico predominante» no significa que no deban dividirse entidades que ejerzan actividades sustanciales en dos o más territorios.

Para el estudio de la distribución de la renta de los hogares que llevaremos a cabo en este trabajo de investigación nos vamos a centrar básicamente en el primero de los tres tipos de análisis existentes en el SEC 2010, el análisis institucional, razón por la cual vamos a detenernos muy brevemente en la definición y características de sus componentes, las unidades y los sectores institucionales.

Para que se pueda hablar de una “unidad institucional” el SEC 2010 exige que se esté en presencia de agentes económicos que posean “autonomía de decisión”⁹ en el ejercicio de la función principal (véase a continuación cuadro 1.1 sobre sectores institucionales) y a las que se les exige además un requisito contable: que dispongan, o tengan capacidad para disponer de un sistema completo de cuentas. En otros términos, éste es un requisito estadístico para poder delimitar en la práctica estas unidades.

Como se recoge en el cuadro 1.1, la definición de los sectores institucionales, esto es, la clasificación de las unidades institucionales en uno u otro grupo, se realiza en principio de acuerdo con dos características fundamentales: su función principal, es decir, el tipo de actuación económica más representativa de cada grupo de unidades; y la categoría de “tipo de productor” (si se dedica a realizar una producción “de mercado” o “no de mercado”)¹⁰ en la que se encuadra. Las unidades institucionales se agrupan en cinco sectores institucionales nacionales mutuamente excluyentes según el tipo de productor de que se trata y de acuerdo con su función y actividad principal, aspectos que se consideran representativos de su comportamiento económico:

- Sociedades no financieras (S.11), caracterizadas por producir bienes y servicios no financieros.
- Instituciones financieras (S.12), dedicadas a la intermediación financiera o a actividades auxiliares de la intermediación financiera.
- Administraciones públicas (S.13), con las funciones tradicionales de este tipo de instituciones como la producción de servicios colectivos y, sobre todo, desde la perspectiva de este trabajo, la redistribución de la renta y la riqueza.
- Hogares (S.14), caracterizados, como función más representativa, por las actividades de consumo final.
- Instituciones sin fines de lucro al servicio de los hogares (S.15), que abarcan un tipo de unidades de creciente expansión en los sistemas económicos actuales y que se caracterizan por suministrar a los hogares fundamentalmente servicios situados fuera del ámbito de las relaciones de mercado (sindicatos, asociaciones religiosas, organizaciones no lucrativas, etc.).

El conjunto de estos cinco sectores constituye el total de la economía nacional. Cada sector está dividido, a su vez, en subsectores. El sistema SEC 2010 permite elaborar una serie completa de cuentas de flujos y de balances para cada sector y subsector, y para el total de la economía. Asimismo, los sectores institucionales mencionados anteriormente son grupos de unidades institucionales residentes.

Las unidades no residentes, es decir, unidades que son residentes de otras economías, pueden interactuar con estos cinco sectores nacionales, y las

⁹ Una unidad institucional goza de autonomía de decisión si se cumplen tres condiciones: es titular de bienes y activos; tiene capacidad para tomar decisiones económicas; y tiene capacidad para contraer pasivos en nombre propio.

¹⁰ Se entiende por producción no de mercado la producción suministrada a otras unidades de forma gratuita o a precios económicamente no significativos, esto es, aquella cuyas ventas no cubren más del 50 por ciento de los costes de producción.

interacciones se presentan entre los cinco sectores nacionales y un sexto sector institucional: el resto del mundo (S.2).

Cuadro 1.1
SECTORES INSTITUCIONALES EN EL SEC 2010

<i>Sector</i>	<i>Tipo de productor</i>	<i>Actividad y función principal</i>
Sociedades no financieras (S.11)	Productor de mercado	Producción de bienes y servicios no financieros de mercado
Instituciones financieras (S.12)	Producto de mercado	Intermediación financiera, incluido el seguro Actividades auxiliares de la intermediación financiera
Administraciones públicas (S.13)	Productor no de mercado	Producción y suministro de bienes y servicios no de mercado para consumo individual y colectivo y realización de operaciones de redistribución de la renta y de la riqueza nacional
Hogares (S.14)		
– como consumidores	– Productor de mercado o	– Consumo
– como empresarios	productor para uso final propio	– Producción de bienes y servicios de mercado y producción de bienes y servicios para uso final propio
Instituciones sin fines de lucro al servicio de los hogares (S.15)	Productor no de mercado privado	Producción y suministro de bienes y servicios no de mercado para consumo individual
Resto del mundo (S.2)	-----	-----

Fuente: SEC 2010

Es importante en este caso aclarar la delimitación precisa del sector Hogares, objeto de estudio a lo largo de esta tesis, de cara a comprender sus relaciones y operaciones de renta: por un lado, hay que precisar que en el sector “Hogares” se incluye una (gran) parte de los empresarios individuales y, por tanto, las funciones de este sector no se limitan al consumo final o la aportación de trabajo al proceso productivo, sino que incluyen actividades de producción. Concretamente, el sector hogares abarca seis subsectores:

- Empleadores (incluidos los trabajadores por cuenta propia).
- Asalariados.
- Perceptores de rentas de la propiedad.
- Perceptores de pensiones.
- Perceptores de rentas procedentes de otras transferencias.
- Otros hogares.

Esta subdivisión, como sugiere la denominación, se establece de acuerdo con el tipo de renta más importante (rentas salariales, rentas de la propiedad, pensiones, etc.) percibido por el hogar en su conjunto, y tiene un indudable interés desde el punto de vista del análisis de la distribución de la renta. Por ello, a continuación, profundizaremos en la definición de cada uno de los subsectores que comprenden el sector hogares.

El subsector «**empleadores (incluidos los trabajadores por cuenta propia)**» está formado por los hogares cuya principal fuente de renta (del hogar en su conjunto) consiste en la renta (mixta) (B.3) percibida por los propietarios de

empresas de los hogares no constituidas en sociedad, que emplean o no personal asalariado, en su actividad de productores de bienes y servicios de mercado, aunque este tipo de renta no suponga más de la mitad de la renta total del hogar.

El subsector **«asalariados»** está compuesto por los hogares cuya principal fuente de renta (del hogar en su conjunto) consiste en la renta procedente de la remuneración de los asalariados (D.1).

Por su parte, el subsector **«perceptores de rentas de la propiedad»** está compuesto por los hogares cuya principal fuente de renta (del hogar en su conjunto) consiste en las rentas de la propiedad (D.4).

El subsector **«perceptores de pensiones»** está compuesto por los hogares cuya principal fuente de renta (del hogar en su conjunto) consiste en la renta procedente de pensiones. Los hogares como pensionistas son aquellos cuya renta procede, principalmente, de pensiones, de jubilación u otras, incluidas las pensiones pagadas por antiguos empleadores.

Y, finalmente, el subsector **«perceptores de rentas procedentes de otras transferencias»** está compuesto por los hogares cuya principal fuente de renta (del hogar en su conjunto) consiste en la renta que procede de otras transferencias corrientes. Las otras transferencias corrientes corresponden a todas aquellas rentas distintas de los salarios, las rentas de la propiedad y las pensiones. Nos estaríamos refiriendo, por ejemplo, a las transferencias periódicas percibidas de otros hogares, indemnizaciones de seguros no vida, premios por loterías, etc.

2.1.2. Las operaciones económicas y el sistema global de cuentas como marco del proceso de distribución de la renta

En lo que se refiere a las variables económicas, dentro de los flujos se recogen las denominadas “operaciones” económicas, las cuales engloban una enorme diversidad de transacciones en el mundo real. En el SEC 2010 se clasifican en cuatro grandes categorías:

- Operaciones de bienes y servicios. Son aquellas que describen el origen (producción interior o importaciones) y el destino (consumos intermedios, consumo final, formación de capital o exportaciones) de los bienes y servicios.
- Operaciones de distribución. Son aquellas que describen cómo se distribuye el valor añadido generado en el proceso productivo entre el factor trabajo, el capital y las administraciones públicas, y la redistribución de la renta y la riqueza (impuestos sobre la renta y el patrimonio, y otras transferencias).
- Operaciones financieras. Son las que describen la adquisición neta de activos financieros o la contracción neta para cada tipo de instrumento financiero. Este tipo de operaciones tienen lugar a menudo como contrapartida de operaciones no financieras, pero también pueden ser operaciones en las que sólo intervienen instrumentos financieros.

- Operaciones no incluidas en los tres grupos anteriores. Comprenden el consumo de capital fijo y las adquisiciones menos las cesiones de activos no financieros no producidos (terrenos, patentes, licencias...).

Para el desarrollo del trabajo que nos ocupa, centraremos nuestra atención en el grupo de las operaciones de distribución, pues es aquí donde se incluyen todos los aspectos de distribución de renta objeto de estudio en este trabajo, desde la distribución primaria entre trabajo y capital (a partir de operaciones como la remuneración de asalariados, el excedente de explotación, los impuestos sobre la producción, ...) hasta la redistribución de las rentas primarias que se realiza entre los distintos agentes de la economía (es el caso de la redistribución de renta impulsada por las administraciones públicas a través de impuestos sobre la renta y el patrimonio, prestaciones y cotizaciones sociales, etc.).

Tal y como encontramos en Cañada (1999), el proceso de distribución de la renta se enmarca en una visión global mucho más amplia: el núcleo central de la Contabilidad Nacional -en su enfoque institucional- es un sistema concatenado de cuentas, de planteamiento similar al de la contabilidad empresarial, en la que se van reflejando de forma ordenada los procesos económicos, y en el que cada cuenta presenta una variable económica fundamental, denominada **saldo contable**¹¹, obtenida como diferencia entre el valor total de los recursos de un lado de una cuenta y el valor total de los empleos del otro lado. Decimos que es un sistema concatenado puesto que las cuentas están ordenadas de forma que el saldo de cada cuenta pasa a convertirse en recurso de la siguiente. Esta sucesión de cuentas se elabora tanto por sectores institucionales como para el total de la economía.

La sucesión completa de las cuentas de las unidades y sectores institucionales comprende una estructura compleja y detallada de operaciones, agrupando las cuentas en tres bloques (cuadro 1.2):

¹¹ Todos los saldos contables pueden calcularse brutos o netos del **consumo de capital fijo**. El consumo de capital fijo representa el montante de los activos fijos consumidos durante el período considerado como resultado del desgaste normal y la obsolescencia previsible, incluida una provisión para las pérdidas de activos fijos como consecuencia de daños accidentales asegurables. Este concepto de consumo de capital fijo debe distinguirse de las amortizaciones aplicadas con fines fiscales o contables que, como veremos más adelante, se utilizan en las encuestas para distinguir entre una variable económica en términos brutos o netos.

**Cuadro 1.2 PRESENTACIÓN SINÓPTICA DE LAS CUENTAS DE LOS SECTORES
INSTITUCIONALES, SALDOS CONTABLES Y PRINCIPALES AGREGADOS EN EL SEC 2010**

	CUENTAS	SALDOS CONTABLES (Y PRINCIPALES AGREGADOS PARA LA ECONOMÍA NACIONAL)
Cuentas corrientes	I. Cuenta de producción	Valor añadido (Producto Interior Bruto/ Producto Interior Neto)
	II. Cuentas de distribución y utilización de la renta	
	II.1 Cuentas de distribución primaria de la renta	
	II.1.1 Cuenta de explotación	Excedente de explotación/ Renta mixta
	II.1.2 Cuenta de asignación de la renta primaria	Saldo de rentas primarias (Renta nacional)
	II.1.2.1 Cuenta de renta empresarial	Renta empresarial
	II.1.2.2 Cuenta de asignación de la renta primaria	Saldo de rentas primarias (Renta nacional)
	II.2 Cuenta de distribución secundaria de la renta	Renta disponible (Renta disponible nacional)
	II.3 Cuenta de redistribución de la renta en especie	Renta disponible ajustada (Renta disponible ajustada nacional)
	II.4 Cuenta de utilización de la renta	Ahorro (Ahorro nacional)
	II.4.1 Cuenta de utilización de la renta disponible	Ahorro (Ahorro nacional)
	II.4.2 Cuenta de utilización de la renta disponible ajustada	Ahorro (Ahorro nacional)
Cuentas de acumulación	III. Cuentas de acumulación	
	III.1 Cuenta de capital	
	III.1.1 Cuenta de variaciones del patrimonio neto debidas al ahorro y a las transferencias de capital	Variaciones del patrimonio neto debidas al ahorro y a las transferencias de capital
	III.1.2 Cuenta de adquisición de activos no financieros	Capacidad/ Necesidad de financiación
	III.2 Cuenta financiera	Capacidad/ Necesidad de financiación
	III.3 Cuenta de otras variaciones de los activos	Variaciones del patrimonio neto debidas al ahorro y a las transferencias de capital
	III.3.1 Cuenta de otras variaciones del volumen de los activos	Variaciones del patrimonio neto debidas a otras variaciones del volumen de los activos
	III.3.2 Cuenta de revalorización	Variaciones del patrimonio neto debidas a ganancias y pérdidas de posesión nominales
	III.3.2.1 Cuenta de ganancias y pérdidas de posesión neutrales	Variaciones del patrimonio neto debidas a ganancias y pérdidas de posesión neutrales
	III.3.2.2 Cuenta de ganancias y pérdidas de posesión reales	Variaciones del patrimonio neto debidas a ganancias y pérdidas de posesión reales
Balances	IV. Balances	
	IV.I Balance de apertura	Patrimonio neto (Patrimonio nacional)
	IV.II Variaciones del Balance	Variaciones del patrimonio neto (Variaciones del patrimonio nacional)
	IV.III Balance de cierre	Patrimonio neto (Patrimonio nacional)

Fuente: SEC 2010

- Las **cuentas corrientes**, que se ocupan de describir una serie de operaciones básicas para la economía que dan lugar a los correspondientes tipos de cuentas: la producción; la distribución de la renta con sus diferentes fases; y la utilización de la renta disponible en las dos grandes categorías macroeconómicas: consumo final y ahorro, factor esencial que nos permite enlazar con las cuentas de acumulación.
- Las **cuentas de acumulación**, en las que se recogen todos aquellos elementos que implican cambios en los niveles y composición del patrimonio de los diferentes sectores (variación del patrimonio neto debidas al ahorro y a las transferencias de capital, adquisiciones de activos no financieros y financieros...). En definitiva, las cuentas de acumulación registran las variaciones del patrimonio neto (que consiste en la diferencia entre los activos y los pasivos). De ellas se obtiene como saldo la capacidad (o necesidad) de financiación.
- Los **balances**, por su parte, presentan los stocks de activos y pasivos al inicio y al final de cada período contable, así como la variación del patrimonio neto total para cada sector y para la economía en su conjunto.

Con el objeto de ofrecer una visión de conjunto, en el cuadro 1.2 se presenta de forma sinóptica la sucesión completa de cuentas de los sectores institucionales, así como sus correspondientes saldos contables y agregados (saldos referidos al total de la economía) en el SEC 2010.

Dado que a lo largo de nuestra investigación nos ocuparemos del estudio de la distribución de la renta del sector hogares, es de vital importancia conocer las diferentes fases y operaciones relacionadas con este proceso en el que será nuestro marco de referencia (la Contabilidad Nacional) así como las cuentas implicadas, integradas dentro del bloque de las cuentas corrientes. La figura 1.1 (véase cuadro 1.3 con el detalle de las cuentas) nos sirve de apoyo para tal fin. En la parte superior derecha de dicha figura aparece la renta obtenida por los hogares en su calidad de productores (que no es otra cosa que el valor añadido) y de la que, una vez descontado el pago a los factores de producción que emplean los hogares (remuneración de asalariados, otros impuestos sobre la producción -y sumando la cuantía de las subvenciones-), aparecerá eventualmente un “excedente” de la actividad que en el caso de los hogares es lo que se denomina “renta mixta”¹². Esta renta mixta, junto con el saldo de las rentas de la propiedad correspondientes a la actividad de producción, dará lugar a la “renta empresarial”¹³.

Sumándole la remuneración de asalariados y las rentas de la propiedad (intereses, dividendos; etc.) no ligadas a la actividad empresarial se tiene el saldo de rentas primarias.

Desde la renta primaria entran en juego los procesos de redistribución 'monetarios' (a través de impuestos sobre renta, cotizaciones, prestaciones...) y se llega a la renta que finalmente está a disposición de las familias para sus actividades de consumo y ahorro, la denominada precisamente **renta disponible de los hogares**. Es decir, por un lado, se añaden a la renta primaria

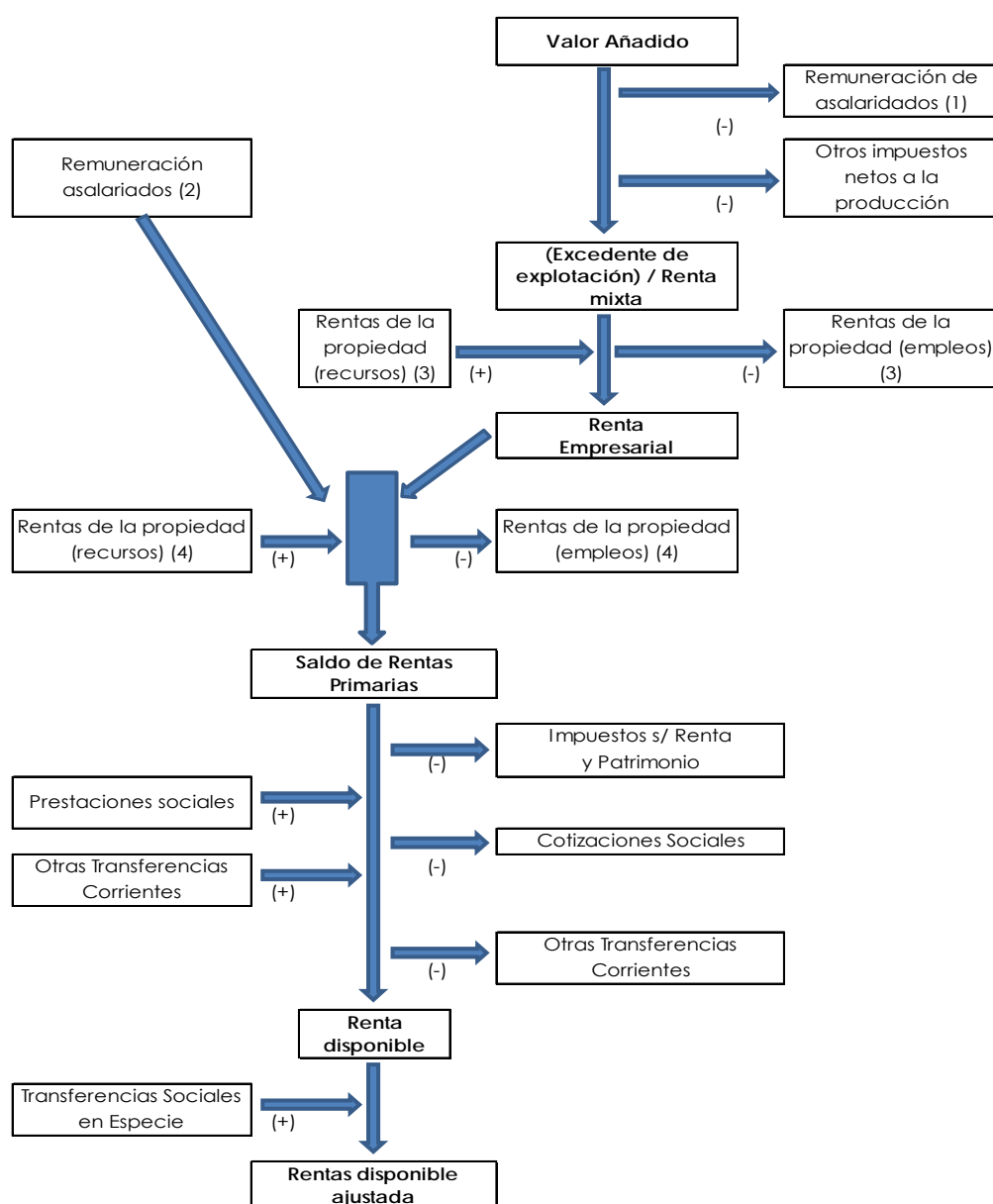
¹² Se ha mantenido el concepto de “excedente” porque se aplica por convenio a las rentas (imputadas) a los hogares que ocupan viviendas en propiedad.

¹³ Como ya veremos, en las encuestas no vamos a contar con este desglose, de modo que tendremos que trabajar con una versión más agregada de las cuentas de distribución de la renta.

recursos como las prestaciones u otras transferencias recibidas, fundamentalmente del sector público, en forma de subsidios y pensiones, detrayéndose por otro aquellas partidas que suponen una disminución de la renta, tales como las cotizaciones sociales o los impuestos directos.

Por último, si a la renta disponible se le añade el saldo de las 'transferencias en especie' se llegaría a la **renta disponible ajustada de los hogares**.

Figura 1.1
LA RENTA DEL SECTOR HOGARES Y SUS COMPONENTES EN EL SEC 2010



(1) Remuneración pagada por los hogares en su calidad de productores.

(2) Remuneración obtenida por los hogares en su calidad de asalariados.

(3) Rentas pagadas/obtenidas por los hogares en su calidad de productores (de mercado).

(4) Rentas pagadas/obtenidas por los hogares no relacionadas con su producción (de mercado).

Fuente: Cañada (1999).

Cuadro 1.3

CUENTAS DE DISTRIBUCIÓN DE LA RENTA PARA EL SECTOR HOGARES EN EL SEC 2010

3.1 Cuentas de distribución primaria de la renta para el sector hogares

a) Cuenta de explotación

<i>Empleos</i>	<i>Recursos</i>
D.1 Remuneración de asalariados	B.1 Valor añadido
D.11 Sueldos y salarios	
D.12 Cotizaciones sociales a cargo de los empleadores	
D.29 Otros impuestos sobre la producción	
D.39 Otras subvenciones a la producción	
B.2 Excedente de explotación/B.3 Renta mixta	

b) Cuenta de renta empresarial

<i>Empleos</i>	<i>Recursos</i>
D.4 Rentas de la propiedad	B.2 Excedente de explotación/B.3 Renta mixta
D.41 Intereses	D.4 Rentas de la propiedad
D.45 Rentas de la tierra	D.41 Intereses
	D.42 Rentas distribuidas de las sociedades

B.4 Renta empresarial

c) Cuenta de asignación de otra renta primaria

<i>Empleos</i>	<i>Recursos</i>
D.4 Rentas de la propiedad	B.2 Excedente de explotación/B.3 Renta mixta
D.41 Intereses	D.1 Remuneración de asalariados
D.45 Rentas de la tierra	D.11 Sueldos y salarios
	D.12 Cotizaciones sociales a cargo de los empleadores
	D.4 Rentas de la propiedad
	D.41 Intereses
	D.42 Rentas distribuidas de las sociedades
	D.44 Otras rentas de inversión
	D.45 Rentas de la tierra

B.5 Saldo de rentas primarias

3.2 Cuentas de distribución secundaria de la renta para el sector hogares

a) Cuenta de distribución secundaria de la renta

<i>Empleos</i>	<i>Recursos</i>
D.5 Impuestos corrientes sobre la renta, el patrimonio, etc	B.5 Saldo de rentas primarias
D.61 Cotizaciones sociales netas	D.5 Impuestos corrientes sobre la renta, el patrimonio, etc
D.62 Prestaciones sociales distintas de las transferencias sociales en especie	D.61 Cotizaciones sociales netas
D.7 Otras transferencias corrientes	D.62 Prestaciones sociales distintas de las transferencias sociales en especie
	D.7 Otras transferencias corrientes

B.6 Renta disponible

b) Cuenta de redistribución de la renta en especie

<i>Empleos</i>	<i>Recursos</i>
	B.6 Renta disponible
	D.63 Transferencias sociales en especie
	D.631 Transferencias sociales en especie: producción no de mercado
	D.632 Transferencias sociales en especie: producción adquirida en el mercado

B.7 Renta disponible ajustada

Fuente: SEC 2010

2.1.3. Las Cuentas Regionales

Las cuentas regionales constituyen una pormenorización regional de las cuentas correspondientes a escala nacional y, en general, utilizan sus mismos conceptos. Las cuentas regionales facilitan un desglose regional de los principales agregados, como el valor añadido bruto por rama de actividad o la renta de los hogares.

Los totales nacionales ocultan las diferencias existentes entre las regiones en lo que respecta a las condiciones y a los resultados económicos. Por lo general, la población y las actividades económicas están distribuidas de forma desigual entre las regiones. Las regiones urbanas suelen estar especializadas en servicios, mientras que la agricultura y las industrias extractivas y manufactureras tienden a estar localizadas en regiones no urbanas. En muchos casos, cuestiones importantes, como la globalización, la innovación, el envejecimiento, la fiscalidad, la pobreza, el desempleo o el medio ambiente, tienen una dimensión económica regional. Por consiguiente, las cuentas regionales son un complemento importante de las cuentas nacionales.

Las cuentas regionales permiten observar las estructuras, los avances y las diferencias económicas regionales. No obstante, debido a determinados problemas conceptuales y de medición, las cuentas regionales son más limitadas que las nacionales en lo que respecta a su alcance y grado de detalle.

2.1.3.1. Conceptos básicos del sistema de cuentas regionales

TERRITORIO REGIONAL

La economía regional de un país es parte del total de la economía de dicho país. Como ya hemos estudiado con anterioridad, el total de la economía se define en términos de unidades institucionales y sectores, y comprende todas las unidades institucionales que tienen un centro de interés económico predominante en el territorio económico de un país. A su vez, el territorio económico de un país se divide en los territorios regionales y el territorio extrarregional.

Según el SEC 2010, el territorio regional consiste en la parte del territorio económico de un país que está directamente asignado a una región, incluidas las zonas francas, fábricas y depósitos aduaneros.

Por su parte, el territorio extrarregional se compone de las partes del territorio económico de un país que no pueden atribuirse directamente a una sola región y está formado por:

a) el espacio aéreo nacional, las aguas territoriales y la plataforma continental situada en aguas internacionales sobre la que el país disfruta de derechos exclusivos;

b) los enclaves territoriales, es decir, los territorios geográficos situados en el resto del mundo y utilizados, en virtud de tratados internacionales o de

acuerdos entre Estados, por las administraciones públicas del país (embajadas, consulados, bases militares, bases científicas, etc.);

c) los yacimientos de petróleo, gas natural, etc. situados en aguas internacionales, fuera de la plataforma continental del país, explotados por unidades residentes.

La nomenclatura de las unidades territoriales estadísticas (NUTS) proporciona un desglose único y uniforme del territorio económico de la Unión Europea. La NUTS es la clasificación territorial para la elaboración de las cuentas regionales.

De acuerdo a estas consideraciones, el equivalente regional del PIB es el Producto Interior Bruto Regional (PIBR). El PIBR se valora a precios de mercado sumando los impuestos menos las subvenciones sobre los productos y las importaciones regionalizados a los valores añadidos por regiones valorados a precios básicos. La suma del PIBR a precios de mercado de todas las regiones, incluido el PIBR del territorio extrarregional, es igual al PIB a precios de mercado.

UNIDADES Y CUENTAS REGIONALES

Como ya hemos señalado con anterioridad, en la economía nacional se distinguen dos tipos de unidades: las unidades institucionales, para analizar los flujos relativos a la renta, el capital y las operaciones financieras, los otros flujos y los balances; y las unidades de actividad económica local (UAE local), que muestran los flujos existentes en el proceso de producción y de utilización de los bienes y servicios.

La UAE local es la parte de una unidad de actividad económica (UAE) que corresponde a una unidad local. La unidad local es una unidad institucional o una parte de unidad institucional que produce bienes y servicios sita en un lugar identificado topográficamente. Cuando una UAE realiza actividades de producción en varias regiones, la información sobre ella se desglosa a fin de obtener las cuentas regionales. Para efectuar dicho desglose, se necesita información sobre la remuneración de los asalariados o, en su defecto, sobre el empleo y sobre la formación bruta de capital fijo. Para las empresas sitas en una localización única y en las cuales la actividad principal represente la parte más importante de su valor añadido, la UAE local coincide con la empresa.

En el contexto de la contabilidad regional es preciso distinguir dos tipos de unidades institucionales, dependiendo del nivel regional de que se trate. En primer lugar, las unidades unirregionales, cuyo centro de interés económico predominante radica en una sola región. Entre las unidades unirregionales figuran los hogares, las sociedades cuyas UAE locales se hallan en la misma región, la mayoría de las administraciones públicas regionales y locales, y parte de las administraciones de seguridad social y de las ISFLSH. Y en segundo lugar, las unidades multirregionales, cuyo centro de interés económico predominante se halla en más de una región. Las sociedades y las ISFLSH son ejemplos de unidades que abarcan varias regiones. Otros ejemplos son las unidades institucionales cuyas actividades se extienden por todo el país, como la administración central y un reducido número de sociedades que ejercen un

monopolio o un cuasimonopolio, como pueden ser los ferrocarriles o las compañías eléctricas nacionales.

Todas las operaciones de las unidades institucionales unirregionales se asignan a la región en que radica su centro de interés económico. En lo que se refiere a los hogares, el centro de interés económico es la región en que residen, no la región en que trabajan. Las demás unidades unirregionales tienen su centro de interés económico en la región en que están situadas.

Algunas de las operaciones de las unidades multirregionales no son, en sentido estricto, conceptualmente regionalizables. Esto sucede con la mayor parte de las operaciones de distribución y las operaciones financieras. Por consiguiente, los saldos contables de las unidades multirregionales, como el ahorro y la capacidad de financiación, no se registran a escala regional.

En base a las consideraciones anteriores, el sistema de cuentas regionales se limita a:

- a) los agregados regionales de las actividades de producción, por ramas de actividad: valor añadido bruto, remuneración de los asalariados, empleo, asalariados y formación bruta de capital fijo;
- b) el producto interior bruto regional (PIBR), que es el resultado de las actividades productivas de las UAE locales residentes en una región;
- c) y las cuentas regionales de los hogares.

MÉTODOS DE REGIONALIZACIÓN

Las cuentas regionales están basadas en las operaciones de las unidades residentes en un territorio regional. De forma general, las cuentas regionales se compilan utilizando los siguientes métodos: métodos ascendentes, métodos descendentes o, una combinación de los anteriores, métodos mixtos.

El método de estimación ascendente supone utilizar la información de las unidades residentes en la región e ir agregando la misma hasta establecer el valor regional del agregado. Los valores regionales tienen que sumarse para obtener el valor nacional correspondiente. Cuando solo se dispone de información a nivel de unidades integradas por varias UAE locales, que realizan actividades distintas y que están situadas en regiones diferentes, en ese caso se utilizan indicadores, como la remuneración de los asalariados y el empleo por región, para obtener las desagregaciones regionales por rama de actividad.

El método descendente consiste en distribuir un total nacional entre las regiones, sin tratar de diferenciar las unidades residentes por región. La cifra nacional se reparte mediante un indicador que se distribuye entre las regiones de la misma forma que la variable que ha de estimarse. La idea subyacente de una unidad residente por región es la que permite alcanzar la cobertura regional del indicador utilizado para asignar regionalmente la variable requerida.

El método ascendente no se presenta, por lo general, en estado puro, por lo que puede considerarse también la posibilidad de utilizar métodos mixtos.

Por ejemplo, las estimaciones regionales de una variable o de un agregado de variables solo se pueden obtener con la ayuda del método ascendente a escala macrorregional y, posteriormente, se utilizará el método descendente para realizar estimaciones a un nivel regional más detallado.

2.1.3.2. Las cuentas regionales de los hogares

Las cuentas regionales de los hogares constituyen una especificación regional de las cuentas correspondientes a escala nacional. Por motivos prácticos, las cuentas se limitan a:

- a) la cuenta de asignación de la renta primaria,
- b) y la cuenta de distribución secundaria de la renta.

En términos generales, estas dos cuentas tienen como objetivo medir la renta primaria y la renta disponible de los hogares residentes en una región determinada (véase cuadro 1.4).

Las cuentas regionales de los hogares se refieren a los hogares residentes en un territorio regional. La suma del número de personas que son miembros de los hogares residentes constituye la población total residente de la región.

Por lo general, las reglas para determinar la residencia de los hogares a escala nacional se aplicarán también a las cuentas regionales de los hogares. Con carácter excepcional, cuando la región de acogida se halle en el mismo país que la región de residencia, los estudiantes y los pacientes de larga duración se tratan como residentes en la región de acogida cuando permanecen en ella más de un año.

Las cuentas de los hogares pueden ampliarse con las cuentas de utilización de la renta. Para ello se tienen que asignar regionalmente las estadísticas de cuentas nacionales relativas al gasto en consumo final de los hogares y al ajuste por la variación de los derechos por pensiones. El saldo contable es el ahorro regional de los hogares.

Para asignar regionalmente el gasto en consumo final de los hogares, es preciso disponer de datos regionales fiables, procedentes, por ejemplo, de una gran encuesta nacional de presupuestos familiares. Sin embargo, en muchos casos no se dispone de ese desglose regional; asimismo, en las cuentas nacionales se utilizan con frecuencia otros datos para estimar el gasto en consumo final de los hogares. En tales circunstancias, resulta más difícil efectuar un desglose regional.

Por su parte, las administraciones públicas pueden desempeñar un papel importante a la hora de prestar a los hogares servicios de educación, asistencia sanitaria y servicios sociales, a través de **transferencias sociales en especie**. El papel de esas transferencias sociales en especie varía considerablemente de un país a otro y puede experimentar cambios importantes en el tiempo. La asignación de esas transferencias sociales en especie por regiones permite obtener el **consumo final efectivo** y la **renta disponible ajustada de los hogares** por regiones. Dado el papel tan relevante que las transferencias sociales en especie tienen en algunos Estados miembros,

la comparación del consumo final efectivo y la renta disponible efectiva de los hogares en los distintos Estados miembros ofrece, en ocasiones, una visión diferente de la que se obtiene comparando el gasto en consumo final y la renta disponible de los hogares.

Cuadro 1.4
CUENTAS REGIONALES DE RENTA DE LOS HOGARES EN EL SEC 2010

	Región			Corrección correspondiente a los flujos regionales sin contrapartida nacional	Total nacional
	a	b	...		
1.4.1 Cuenta de asignación de la renta primaria de los hogares					
<i>Recursos</i>					
B.2/B.3 Excedente de explotación neta/renta mixta neta					
D.1 Remuneración de los asalariados					
D.4 Rentas de la propiedad a cobrar				Menos las rentas de la propiedad intrarregionales relacionadas con B.2/B.3	
<i>Empleos</i>					
D.4 Rentas de la propiedad a pagar				Menos las rentas de la propiedad intrarregionales relacionadas con B.2/B.3	
B.5 Renta primaria (saldo contable)					
1.4.2 Cuenta de distribución secundaria de la renta de los hogares					
<i>Recursos</i>					
B.5 Renta primaria					
D. 62 Prestaciones sociales distintas de las transferencias sociales en especie					
D. 7 Otras transferencias corrientes a cobrar					
<i>Empleos</i>					
D.5 Impuestos corrientes sobre la renta, el patrimonio, etc.					
D.61 Cotizaciones sociales netas					
D.7 Otras transferencias corrientes a pagar					
B.6 Renta disponible (saldo contable)					
Información complementaria					
Población (número de habitantes)					
Renta primaria per cápita					
Renta disponible per cápita					

Fuente: SEC 2010

Finalmente, es preciso destacar respecto a la CRE que, el menor grado de detalle de su información, debido a los problemas conceptuales y de medición mencionados en este apartado, unido a la imposibilidad de desagregar por región los datos de determinadas encuestas que serán de vital importancia en este estudio (como la EFF, que estudiaremos más adelante) nos llevarán a descartar esta fuente como marco contable de referencia y, por tanto, a utilizar CNE como base de la comparación.

3. Sobre las fuentes estadísticas a nivel individual

En el siguiente apartado, nos detendremos en la descripción y delimitación de las fuentes estadísticas que nos facilitan información sobre la renta de los hogares españoles a nivel de microdato, es decir, las encuestas a hogares, así como del período de tiempo para el que están disponibles y las limitaciones que presentan.

En concreto, y tal como anticipábamos en la Introducción, estudiaremos, en primer lugar, la Encuesta de Condiciones de Vida y la Encuesta de

Presupuestos Familiares, elaboradas por el INE, que se caracterizarán como veremos en el detalle de la información económica que proporcionan respecto a los ingresos y gastos de los hogares, respectivamente; a continuación, pasaremos a describir la Encuesta Financiera de las Familias, una operación estadística del Banco de España acerca de la riqueza y sus componentes en la unidad familiar; después nos detendremos en las bases de microdatos del Instituto de Estudios Fiscales a nivel de hogar, compuestas por la muestra de declarantes y no obligados no declarantes del IRPF y el panel de declarantes del IRPF; y, por último, analizaremos la Muestra Continua de Vidas Laborales, a cargo del Instituto Nacional de la Seguridad Social.

3.1. La Encuesta de Condiciones de Vida

La **Encuesta de Condiciones de Vida (ECV)**, en terminología inglesa *European Statistics on Income and Living Conditions (EU-SILC)*, es una operación estadística anual dirigida a hogares que se realiza en todos los países de la Unión Europea. Se trata de una estadística armonizada que está respaldada por el Reglamento (CE) N° 1177/2003 del Parlamento Europeo y del Consejo de 16 de junio de 2003 relativo a las estadísticas comunitarias sobre la renta y las condiciones de vida.

La ECV está disponible desde el año 2004 y se realiza durante la primavera de cada año, normalmente entre marzo y junio. Anteriormente se llevó a cabo la que puede considerarse su predecesora, el Panel de Hogares de la Unión Europea (PHOGUE), realizado durante el periodo 1994-2001. Las características y objetivos de ambas encuestas son similares, aportando el PHOGUE una experiencia muy valiosa para la nueva encuesta.

El objetivo general de la ECV es proporcionar información sobre la renta, el nivel y composición de la pobreza y la exclusión social en España y permitir la realización de comparaciones con otros países de la Unión Europea. Para ello, la ECV tiene el diseño de un panel rotante, es decir, en el que las personas entrevistadas colaboran cuatro años seguidos, incluyendo, de esta forma, información longitudinal y transversal a nivel nacional y europeo.

Para la realización de la ECV se toma una muestra de unas 16.000 viviendas distribuidas en 2000 secciones censales de todo el territorio nacional. El método de recogida de datos es el de entrevista personal a los miembros de cada uno de los hogares incluidos en la muestra.

En la ECV se recogen los ingresos del año natural anterior a la entrevista. Además, se recogen otras muchas preguntas sobre condiciones de vida que se refieren al momento de la entrevista. Así, por ejemplo, las variables sobre ingresos de la encuesta de 2012 se refieren al año 2011 mientras que las demás preguntas se refieren a 2012.

Entre sus utilidades, la encuesta destaca por constituir un elemento de referencia fundamental para el seguimiento del Plan Nacional de Acción para la Inclusión Social.

Más concretamente, la ECV está orientada a proporcionar informaciones comparables y armonizadas sobre los siguientes aspectos del nivel y condiciones de vida y de la cohesión social:

1. Ingresos de los hogares privados y en general sobre su situación económica (la información sobre ingresos está referida al año anterior al de la encuesta).
2. Pobreza, carencias, protección social e igualdad de trato.
3. Empleo y actividad.
4. Jubilaciones, pensiones y situación socioeconómica de las personas mayores.
5. Vivienda y costes asociados a la misma.
6. Desarrollo regional.
7. Nivel de formación, salud y efectos de ambos sobre la condición socioeconómica.

En lo que respecta a los ingresos que se recogen en la ECV, se encuentran tanto los ingresos monetarios como los no monetarios. El objetivo consiste en llegar a obtener la **renta disponible del hogar**, es decir, la renta que resulta una vez deducidas las cantidades satisfechas en concepto de impuestos, cotizaciones a la Seguridad Social y otros pagos asimilados (Entidades de Previsión Social, Mutualidades obligatorias y Derechos Pasivos). Este concepto, al que nos referiremos simplemente como renta, es el término más idóneo para evaluar los niveles de renta de un hogar y el que utilizaremos a lo largo de este estudio.

Concretamente, en la ECV los ingresos se clasifican en los siguientes tipos atendiendo a su origen:

- **Renta del asalariado**, entendida como la remuneración total, monetaria o no monetaria, que debe pagar un empleador a su asalariado como contrapartida del trabajo realizado por éste durante el período de referencia de los ingresos.

La renta del asalariado se desglosa en:

- Renta bruta monetaria o cuasimonetaria del asalariado.
- Renta bruta no monetaria del asalariado.
- Cotizaciones sociales a cargo del empleador.

Estos tres componentes de la renta del asalariado se recogen por separado en la ECV para cada uno de los miembros del hogar.

- **Renta de los trabajadores por cuenta propia**, que comprende los ingresos percibidos durante el período de referencia de los ingresos por las personas, para sí o en relación con los miembros de su familia, como resultado del ejercicio presente o pasado de una actividad por cuenta propia.

La renta de los trabajadores por cuenta propia se desglosa en:

- Beneficios o pérdidas monetarios brutos de trabajadores por cuenta propia (incluidos derechos de propiedad intelectual), cuya información está disponible para cada miembro del hogar.
 - Valor de los bienes producidos para autoconsumo, que se recoge a nivel de hogar.
- **Alquiler imputado**, que se refiere al valor que se imputará a todos los hogares que no declaren pagar un alquiler completo, ya sea porque son ocupantes propietarios, porque ocupan una vivienda alquilada a un precio inferior al de mercado o porque ocupan una vivienda a título gratuito.

El alquiler imputado sólo se calculará en el caso de las viviendas (y demás locales asociados, como garajes) utilizadas por los hogares como residencia principal. Por tanto, en la encuesta esta variable se recoge para cada hogar.

- **Rentas de la propiedad**, que comprende los ingresos percibidos durante el período de referencia de los ingresos por el propietario de un activo financiero o un activo material no producido, como contrapartida por suministrar fondos a otra unidad institucional o por poner a disposición de ésta el activo material no producido, una vez deducidos los gastos soportados.

Las rentas de la propiedad se desglosan en:

- Intereses.
- Dividendos.
- Ganancias de inversiones de capital en empresas no constituidas en sociedad.
- Renta procedente del alquiler de una propiedad o terreno.

Todas estas rentas se recogen a nivel de hogar, las tres primeras conjuntamente en una sola variable.

- **Transferencias corrientes percibidas**, entre las que se encuentran:
- **Prestaciones sociales**, que se definen como las transferencias corrientes percibidas por los hogares durante el período de referencia de los ingresos con objeto de aligerarlos de la carga financiera que suponen determinados riesgos o necesidades, realizadas a través de sistemas organizados colectivamente o por entidades estatales e instituciones sin fines de lucro al servicio de los hogares (ISFLSH).

Las prestaciones sociales se desglosan en:

- Ayudas por familia/hijos.
- Ayudas para vivienda.
- Prestaciones por desempleo.
- Prestaciones por vejez.
- Prestaciones por supervivencia.
- Prestaciones por enfermedad.
- Prestaciones por invalidez.
- Ayudas para estudios.
- Exclusión social no clasificada en otro apartado.

De todas ellas tenemos información disponible para cada uno de los miembros del hogar salvo para las dos primeras, ayudas por familia/hijos y ayudas para vivienda, que por su naturaleza se recogen a nivel de hogar.

- **Transferencias periódicas monetarias percibidas de otros hogares**, que se refieren a cantidades periódicas monetarias percibidas durante el período de referencia de los ingresos procedentes de otros hogares o personas. La información de este tipo de renta está disponible en la encuesta para cada hogar.
- **Otras rentas percibidas: Rentas percibidas por los menores de 16 años.** La renta percibida por los menores de 16 años incluye la renta bruta percibida por todos los miembros del hogar menores de 16 años durante el período de referencia de los ingresos. Esta variable está disponible en la ECV a nivel de hogar.
- **Pago de intereses: Intereses pagados de préstamos hipotecarios**, que se refieren al importe bruto total de los intereses hipotecarios, antes de deducir cualquier reducción fiscal, por la vivienda principal del hogar durante el período de referencia de los ingresos. Al igual que la anterior, esta variable se recoge también para el hogar.
- **Transferencias corrientes pagadas**, las cuales se desglosan en:
 - Impuesto sobre la renta y cotizaciones sociales.
 - Impuesto sobre el patrimonio.
 - Cotizaciones sociales a cargo del empleador.
 - Transferencias periódicas monetarias pagadas a otros hogares.

A excepción de las Cotizaciones sociales a cargo del empleador, estas transferencias corrientes pagadas se investigan para el hogar en su conjunto.

Una vez descritos brevemente los diferentes tipos de renta que se recogen en la ECV, y basándonos en el amplio grado de especialización y detalle que esta encuesta proporciona sobre los ingresos de los hogares, en el ámbito de este trabajo de investigación, vamos a considerar a la ECV como la principal fuente de microdatos para el análisis de la distribución personal de la renta.

Aunque como hemos dicho, esta fuente va a ser la más importante dentro de las fuentes de microdatos que utilizaremos, y a pesar de los continuos esfuerzos para obtener mejoras en la calidad de sus datos, la ECV presenta una serie de limitaciones relativas a la población elegida para la evaluación de la renta, unidades de análisis, fiabilidad de los datos y a la periodicidad de los mismos.

En cuanto a la población elegida para la evaluación de la renta, hay que decir que, aunque toda la población es potencialmente incluíble en el muestreo, la ECV no recoge información de los individuos que carecen de domicilio o que están en situación de residencia no legalizada.

La población objeto de investigación (población objetivo o unidades de análisis) es la del conjunto de hogares privados que residen en viviendas familiares principales y el conjunto de personas, miembros del hogar, de dichos hogares. Aunque las personas de todas las edades forman parte de la

población objetivo no todas las personas son investigadas exhaustivamente (en el sentido que se recoge información individualizada), ya que sólo son seleccionables para una investigación exhaustiva las personas, miembros del hogar, con 16 o más años el 31 de diciembre del año anterior al de entrevista.

En relación a la fiabilidad de los datos, habrá que tener en cuenta la posible subestimación de los ingresos en las encuestas a hogares y la negativa, incapacidad o imposibilidad de los hogares y/o personas a cumplimentar la encuesta, ya sea total o parcialmente (falta de respuesta). En la literatura económica es frecuente encontrar numerosos estudios como los de Alcaide Inchausti, A. y J. (1974), (1977) y (1983), Pena *et al.* (1996), Prieto y Pena (2000), Andrés y Mercader (2001), etc. que han venido evidenciando a lo largo del tiempo estas discrepancias entre ciertas fuentes de información concretas (Encuestas Básicas de Presupuestos Familiares, Panel de Hogares de la Unión Europea, Contabilidad Nacional...) y para distintos períodos (1964, 1967, 1970, 1973, 1981, 1991, 2000...).

Con la finalidad de minimizar el impacto de estos errores ajenos al muestreo sobre la calidad de los datos obtenidos en las encuestas, en la ECV de 2013 el INE adoptó una nueva metodología en la producción de datos relativos a los ingresos del hogar basada en el uso de ficheros administrativos. El acceso a registros administrativos supone una mejora de la calidad de los datos y en la eficiencia en el método de recogida. Así, en la ECV Base 2013 los datos relativos a los ingresos del hogar se han elaborado mediante una metodología mixta combinando la información proporcionada por el informante con los registros administrativos de la Agencia Estatal de Administración Tributaria, la Seguridad Social, la Hacienda Tributaria de Navarra y la Diputación Foral de Bizkaia.

Por último, otro de los puntos a tener en cuenta para trabajar con esta encuesta será la restricción a la que nos enfrentamos respecto al período de tiempo que estamos interesados en estudiar, el cual dependerá de la realización o no de la encuesta durante el mismo. Las ECV disponibles en el momento de realizar nuestro estudio son las correspondientes al período de ingresos comprendido entre los años 2003 y 2011 para las ECV en base 2004, y desde 2007 hasta 2014 para las ECV en base 2013.

3.2. La Encuesta de Presupuestos Familiares

Otra de las encuestas considerada de gran relevancia debido a la información económica que proporciona, es la **Encuesta de Presupuestos Familiares (EPF)**, una de las encuestas más antiguas de las que realiza el Instituto Nacional de Estadística (INE). La actual EPF, iniciada en enero de 2006, suministra información anual sobre la naturaleza y destino de los gastos de consumo, así como sobre diversas características relativas a las condiciones de vida, entre las que figura la renta de los hogares. Dicha encuesta, sustituye a la Encuesta Continua de Presupuestos Familiares (ECPF) base 1997, que con periodicidad trimestral se realizó desde 1997 hasta 2005.

La EPF obtiene su información mediante entrevista personal a cerca de 24.000 hogares seleccionados aleatoriamente, que colaboran durante dos

semanas consecutivas en cada uno de los dos años que permanecen en la muestra.

Entre los objetivos prioritarios de la EPF se encuentran: proporcionar estimaciones del agregado gasto de consumo anual de los hogares para el conjunto nacional y para las comunidades autónomas, así como su clasificación según diversas variables del hogar; estimar el cambio interanual del agregado gasto de consumo para el conjunto nacional y para las comunidades autónomas; y la estimación del consumo en cantidades físicas de determinados bienes alimenticios y fuentes de energía para el conjunto nacional. Además, dentro de los objetivos principales también destacan por su importancia otros dos relacionados con necesidades concretas de los principales usuarios de la encuesta: la estimación del gasto como instrumento para la obtención del consumo privado en Contabilidad Nacional y la estimación de la estructura de ponderaciones a partir del gasto necesaria para el cálculo del IPC.

Los gastos de consumo recogidos en la EPF se refieren tanto al flujo monetario que destina el hogar al pago de determinados bienes y servicios de consumo final, como al valor de los bienes percibidos en concepto de autoconsumo, autosuministro, salario en especie, comidas gratuitas o bonificadas y alquiler imputado a la vivienda en la que reside el hogar (cuando es propietario de la misma o la tiene cedida por otros hogares o instituciones). Los gastos se registran en el momento de adquisición, independientemente de que el pago sea al contado o a plazos.

La clasificación utilizada para recoger los gastos de consumo realizado por los hogares es la COICOP, adaptación nacional de la clasificación internacional utilizada en Eurostat para las encuestas de presupuestos (COICOP/HBS) -Classification Of Individual Consumption by Purpose-, y que está estructurada en doce grandes grupos:

1. Alimentos y bebidas no alcohólicas
2. Bebidas alcohólicas, tabaco y narcóticos
3. Artículos de vestir y calzado
4. Vivienda, agua, electricidad, gas y otros combustibles
5. Mobiliario, equipamiento del hogar y gastos corrientes de conservación de la vivienda
6. Salud
7. Transportes
8. Comunicaciones
9. Ocio, espectáculos y cultura
10. Enseñanza
11. Hoteles, cafés y restaurantes
12. Otros bienes y servicios

El nivel de desagregación funcional del que disponemos en la publicación de resultados varía según el grado de desagregación geográfica al que estén referidas las estimaciones.

Por lo que respecta a la información sobre rentas, en la EPF además de solicitar el ingreso del hogar en su conjunto se recoge información sobre los **ingresos individuales**, esto es, se pregunta sobre el número de miembros del hogar que son perceptores de ingresos y su ingreso puntual, de forma que si no se facilita un valor puntual se solicita un intervalo en el que estén comprendidos esos ingresos individuales.

Para los **ingresos del hogar** se tienen en cuenta los ingresos monetarios percibidos regularmente por los miembros del hogar cualquiera que sea su origen. Así, se considera que los miembros del hogar perciben ingresos cuando están trabajando o tienen un trabajo por el que perciben una remuneración, o perciben alguna pensión, subsidios, rentas de capital o transferencias. Los ingresos mensuales en la actualidad se refieren a los ingresos regulares percibidos en el último mes de calendario, teniendo en cuenta los prorrateos de los ingresos extraordinarios y otros ingresos que se reciban de forma periódica, aunque no mensualmente. En concreto, la EPF recoge información sobre los siguientes tipos de fuente de ingresos:

- **Ingresos de trabajo por cuenta propia** (ingresos netos, es decir, ingresos brutos menos los gastos deducibles y las retenciones a cuenta y pagos fraccionados).
- **Ingresos de trabajo por cuenta ajena** (sueldos, salarios y antigüedad, horas extraordinarias, guardias, pagas extraordinarias, complementos de productividad, específicos y de destino, gratificaciones).
- **Ingresos por pensiones contributivas y no contributivas** (vejez, jubilación o retiro, invalidez, viudedad, orfandad o en favor de familiares).
- **Subsidios y prestaciones de desempleo, formación profesional o de fomento de empleo.**
- **Otros subsidios y prestaciones sociales regulares** (enfermedad, mantenimiento del salario, ayudas a la familia por natalidad, maternidad, protección familiar, becas).
- **Rentas de la propiedad y del capital** (alquileres, dividendos, intereses, rendimientos).
- **Otros ingresos regulares** (donaciones de instituciones, transferencias de otros hogares, remesas de emigrantes y otros ingresos regulares distintos a las prestaciones sociales).

Además, la EPF dispone de dos variables más relacionadas con los ingresos:

- la variable **principal fuente de ingresos del hogar**, que es la que resulta mayor después de agregar los ingresos de los distintos miembros del hogar por tipo;
- y los **ingresos mensuales netos del hogar**, que comprenden el conjunto de los ingresos percibidos regularmente por el hogar y sus miembros perceptores de ingresos individuales, cualquiera que sea su origen, una vez descontadas las cotizaciones a la Seguridad Social y otros pagos asimilados (Entidad de Previsión Social, Mutualidades Obligatorias y Derechos Pasivos), así como las cantidades satisfechas en concepto de impuestos. En el caso de trabajo por cuenta ajena, deben tomarse en

consideración exclusivamente las retenciones a cuenta del Impuesto del Rendimiento de las personas Físicas (IRPF), mientras que, en el caso de trabajo por cuenta propia, deben tomarse en consideración los impuestos retenidos a cuenta y los pagos fraccionados del IRPF.

A pesar de contener información tanto de ingresos como de gastos, como más adelante veremos, aprovechando el elevado nivel de detalle con el que se recogen los gastos en la EPF, la utilidad de esta encuesta en nuestro estudio será la de poder relacionar el nivel de ingresos de un hogar en base a sus gastos totales declarados, siendo el período de referencia de los mismos el año de la entrevista.

Al igual que hiciéramos en la presentación de la ECV, a la hora de analizar los datos de la EPF, no debemos pasar por alto las limitaciones que presenta. Algunas de ellas, como las relativas a la población a la que se refieren las variables económicas coinciden con las ya mencionadas en relación a la ECV, ya que este concepto coincide en ambas encuestas.

Asimismo, en cuanto a la fiabilidad de los datos, tal y como numerosos estudios relativos a las encuestas a hogares han puesto de manifiesto¹⁴, debemos saber que cuando se recogen simultáneamente ingresos y gastos, los datos de gastos suelen ser más fiables y mayores que los ingresos declarados, permitiendo este hecho emplear los gastos para proceder a la corrección de los ingresos, al alza, con la finalidad de mejorar su consistencia. Aspectos relativos a la naturaleza de la EPF tales como la posible existencia de falta de respuesta total o parcial también tendremos que tener en cuenta, si bien cuando la falta de respuesta afecta a una variable cuantitativa se recurre a técnicas de imputación de los valores faltantes.

Finalmente, cabe destacar que para la selección de nuestro periodo de investigación debemos tener presente que a fecha de ejecución de esta tesis se encuentra disponible la información de las EPF base 2006 pertenecientes al período 2006-2015.

3.3. La Encuesta Financiera de las Familias

Para relacionar las rentas, los activos, las deudas y el gasto de cada unidad familiar, nos vamos a servir de la única encuesta disponible en España que nos va a permitir llevarlo a cabo, la Encuesta Financiera de las Familias (EFF). Esta encuesta, elaborada por el Banco de España con la colaboración del INE y la Agencia Estatal de Administración Tributaria (AT), se realizó por primera vez¹⁵ en 2002, posteriormente en 2005, 2008 y 2011, hasta la última edición disponible, la EFF 2014, cuya referencia temporal es diciembre de dicho año¹⁶, y sirve para dar continuidad a esta fuente estadística.

¹⁴ Attanasio, O., Battistin, E. and Leicester, A. (2006), Holly Sutherland H., Taylor, R. and Gomulka, J. (2011), Brewer, M. and O'Dea C. (2012), etc.

¹⁵ En Bover (2004) se realiza una descripción detallada de la primera ola de la EFF.

¹⁶ Dada la utilidad de la información proporcionada por una encuesta de las características de la EFF, el sistema de bancos centrales del área del euro realiza desde 2011 la *Household Finance and Consumption Survey* (HFCS), una encuesta armonizada sobre la situación financiera y el consumo de los hogares que sigue una metodología similar a la de la EFF. En la primera ola, se entrevistaron 62.000 hogares en quince países de

La EFF más reciente, la EFF 2014, actualiza información sobre las finanzas de los hogares y, de esta manera, permite una valoración de los cambios ocurridos en la situación financiera de las familias españolas durante el período transcurrido entre finales de 2011 y finales de 2014, a la vez que ofrece una imagen representativa más actualizada de la composición y la distribución de los activos y de las deudas de los hogares. Además, la disponibilidad de datos para un período considerablemente amplio como 2002-2014 permite analizar dos fases completas del ciclo económico, las cuales han tenido un fuerte impacto en la situación patrimonial de los hogares españoles.

Una característica relevante de las cinco ediciones de la EFF es que sus muestras contienen un número elevado de hogares con un alto nivel de riqueza. La distribución de la riqueza es muy asimétrica y, además, algunas clases de activos solo están en poder de una pequeña fracción de la población. Por consiguiente, a menos que se puedan recoger muestras de tamaño muy elevado, es importante incorporar, como se hace en la EFF, un sobremuestreo de los hogares de mayor nivel de riqueza para contar con una muestra que no solo sea representativa del conjunto de la población, sino también de la riqueza agregada de la economía y que facilite el estudio del comportamiento financiero en el tramo superior de su distribución.

Una importante característica adicional de la EFF es que a partir de la segunda edición se ha entrevistado a una parte de los hogares que colaboraron en ediciones anteriores. Así, la combinación de las muestras de las distintas ediciones permite observar un subconjunto de hogares en varios momentos del tiempo y, en algunos casos, durante un período de casi diez años¹⁷, lo que amplía las posibilidades de análisis sobre el comportamiento de la renta, la riqueza y el consumo a lo largo del ciclo vital de los hogares, así como para explorar las transiciones o la movilidad de los hogares a lo largo de las distribuciones de las variables objeto de estudio. Estas características, se han obtenido gracias a un sistema de colaboración con el INE y la AT.

El cuestionario de la EFF se divide en nueve secciones principales, que reflejan las variables objeto de estudio, y que son las siguientes:

1. Características demográficas
2. Activos reales y deudas asociadas
3. Otras deudas
4. Activos financieros
5. Seguros y pensiones
6. Situación laboral e ingresos relacionados
7. Rentas no relacionadas con la actividad laboral percibidas en el año natural anterior (2013, para la EFF 2014)
8. Uso de instrumentos de pago

la zona del euro. En la segunda ola, cuyos resultados principales han sido publicados recientemente, se proporciona información para una muestra de 84.000 hogares en un total de 18 países de la zona del euro, y en Hungría y Polonia. En el sitio web del BCE, https://www.ecb.europa.eu/pub/economic-research/research-networks/html/researcher_hfcn.en.html, se puede acceder a los datos micro y a los documentos con los principales resultados de ambas olas.

¹⁷ En el diseño de la muestra de la EFF 2014 no se incluyeron los hogares entrevistados en la EFF 2002.

9. Consumo y ahorro

Las preguntas relativas a los activos y a las deudas se refieren al conjunto del hogar, mientras que las relativas a la situación laboral y a los ingresos relacionados se realizan para cada uno de los miembros del hogar mayores de 16 años. La mayoría de la información se refiere al momento de la entrevista, aunque también se recoge información sobre todas las rentas antes de impuestos referidas al año natural anterior a la encuesta, es decir, para la EFF 2014, a 2013. La recogida de esta información se llevó a cabo mediante entrevistas personales a los hogares realizadas por entrevistadores con formación específica y con la ayuda de ordenador, debido a la complejidad y extensión del cuestionario.

En cuanto al diseño muestral de la EFF, en la de 2014, al igual que en ediciones anteriores, tiene dos objetivos principales. Por un lado, mantener las mismas características de calidad y, en particular, proporcionar una muestra representativa de la población con sobremuestreo de riqueza. Por otro lado, se pretende que parte de la muestra sea longitudinal, es decir, que incluya hogares que colaboraron en realizaciones anteriores de la EFF. En relación con esta dimensión longitudinal, en esta quinta edición de la EFF se ha iniciado un esquema de rotación que limita el número máximo de ediciones en que un hogar permanece en la encuesta. En concreto, en la muestra de la EFF 2014 no se incluye a hogares entrevistados en la EFF 2002. De esta manera se introduce una estructura de subpaneles enlazados que facilita compaginar la representatividad de la actividad económica en cada momento con el componente longitudinal. Para conciliar ambos objetivos, se diseñó una muestra de refresco que complementa el componente panel (hasta un tamaño muestral total de 8.335 hogares) y asegurar que, al ser utilizada junto con la parte longitudinal, la muestra global cumple los requisitos de representatividad y sobremuestreo. Esta muestra se ha logrado gracias a la colaboración del INE y de la AT, a través de un complejo mecanismo de coordinación que permite respetar en todo momento estrictos compromisos de confidencialidad y anonimato en la asignación del nivel de riqueza del hogar, para lograr los objetivos establecidos por el diseño muestral. En concreto, la AT confecciona una muestra aleatoria por estratos de riqueza, adicional a la muestra longitudinal, a partir del Padrón Continuo proporcionado por el INE, siguiendo las directrices del diseño muestral elaborado por este último. Esto asegura la representatividad de la información obtenida, a la vez que se logra información precisa sobre el comportamiento del segmento de hogares con mayores recursos e incorpora un componente longitudinal.

Como en cualquier encuesta, para obtener magnitudes representativas de la población, se utilizan factores de elevación, es decir, la frecuencia con la que los hogares de la muestra se encuentran en la población total de hogares. Para calcular estos factores se consideran, principalmente, las características del diseño muestral, pero se realizan ajustes diversos, en particular para recoger la diferente tasa de no respuesta por niveles de renta y patrimonio.

Al igual que destacábamos en las dos encuestas estudiadas en los apartados 1.3.1 y 1.3.2, la EFF también adolece de problemas de falta de respuesta. Cualquier análisis basado exclusivamente en los casos con los cuestionarios totalmente cumplimentados podría dar lugar a importantes

sesgos en los resultados. La solución adoptada por el Banco de España ha sido la elaboración de imputaciones de los valores no observados para facilitar el análisis de los datos. Dichas imputaciones se basan en técnicas estadísticas avanzadas, obteniéndose cinco estimaciones para cada valor no observado, a fin de tener en cuenta la incertidumbre asociada a la imputación. Los estadísticos objeto de estudio se obtienen combinando la información de estas imputaciones múltiples.

Finalmente, corresponde hablar de las ventajas que presenta esta encuesta, que se pueden resumir en las siguientes:

- Permite obtener información en diferentes momentos del tiempo (2002, 2005, 2008, 2011 y 2014) sobre la posición relativa de los distintos tipos de hogares estudiados, lo que nos permite valorar los cambios ocurridos a lo largo del tiempo.
- La información se completa con datos relativos a su riqueza neta y a su consumo, lo que permite contrastar la situación en términos de renta, riqueza y consumo para la misma muestra de hogares. De hecho, esta encuesta, constituye la única fuente estadística en España que permite relacionar renta, activos y gasto de cada unidad familiar.
- Otra importante característica de esta encuesta es que presenta una parte de la muestra que permanece a lo largo del tiempo, lo que permite observar la evolución de un subconjunto de hogares a lo largo del tiempo y amplía considerablemente las posibilidades de análisis sobre la naturaleza de los cambios observados.

Por su parte, entre las principales limitaciones de la EFF, se encuentran:

- Aunque, al igual que las ECV, las EFF recogen información personal y agregada para el hogar sobre los ingresos de sus componentes, el tamaño muestra de las EFF es bastante menor que el de las ECV (aproximadamente la mitad).
- Los datos obtenidos y analizados se obtienen de la información que suministra el encuestado, de forma que si éste oculta o no ofrece toda la información relevante, al igual que ocurre en las ECV, esa laguna no queda cubierta.
- A diferencia de la ECV y de la EPF, la EFF no proporciona datos desagregados por región, restricción que nos limitará a la hora de valorar la elaboración de nuestro estudio a nivel regional.

3.4. Fuentes fiscales

Del conjunto de fuentes estadísticas tributarias a nivel de microdato destacan, para el caso de los hogares, la Muestra del Impuesto sobre la Renta de las Personas Físicas y el Panel de declarantes de IRPF.

3.4.1. La Muestra del Impuesto sobre la Renta de las Personas Físicas

El Instituto de Estudios Fiscales (IEF), en colaboración con la AT, pone a disposición de los investigadores, con periodicidad anual, la muestra de microdatos correspondiente a las declaraciones del Impuesto sobre la Renta de las Personas Físicas (IRPF) en el territorio de régimen fiscal común (sin incluir Comunidad Foral de Navarra y País Vasco) y la muestra de perceptores no obligados a declarar. Hasta la fecha¹⁸, el período para el que la muestra de IRPF está disponible es el correspondiente a 2002-2013.

La **muestra de declarantes** se compone de las declaraciones presentadas anualmente del IRPF en el Territorio de Régimen Fiscal Común. Estas unidades de muestreo se seleccionan de entre el conjunto de declaraciones cuyo documento de ingreso o devolución es el Modelo 100 (incluye declaraciones ordinarias y simplificadas tanto individuales como conjuntas).

Por su parte, aquellos perceptores pertenecientes al Territorio de Régimen Fiscal Común, cuya suma de retribuciones se encuentra entre 10.000 y 22.000 euros brutos anuales, forman parte de la **muestra de no declarantes**, siendo en este caso las unidades de muestreo las declaraciones que realizan los pagadores mediante el modelo 190, que es el modelo por el que las empresas, autónomos y demás obligados tributarios, liquidan las retenciones derivadas (principalmente de las nóminas de los trabajadores y de las retenciones en factura a profesionales).

La muestra de declarantes y de no obligados-no declarantes presenta una serie de ventajas tales como su gran representatividad debida a la utilización de muestreo estratificado, la ausencia de problemas de infrarrepresentación y falta de respuesta, una alta precisión debida al origen fiscal de los datos, y la elevada representatividad poblacional debido a la inclusión de la muestra de no obligados-no declarantes, formada por aquellas personas que no han presentado la declaración del IRPF pero estaban sometidos a retenciones.

Pese a estas ventajas que se derivan de su amplitud y de la calidad de los datos, la muestra IEF-AT presenta también varias limitaciones, entre las que se encuentran:

- La unidad de análisis es la declaración de IRPF, y la muestra no proporciona ninguna información que permita identificar a los cónyuges pertenecientes a un hogar que hayan declarado individualmente, ni separar las rentas de aquellos que lo han hecho conjuntamente. Este hecho no nos permite reconocer a contribuyentes que convivan en el mismo hogar, de forma que los estudios sobre desigualdad sólo pueden realizarse en función de las unidades definidas en la muestra, es decir, individuos o unidades familiares fiscales.
- No poseemos información de todo el territorio nacional, únicamente del territorio fiscal común que no incluye la Comunidad Foral de Navarra y el País Vasco.
- Falta de cualquier información extra-fiscal no necesaria para la liquidación del impuesto correspondiente.

¹⁸ Última entrada a 13 de abril de 2017 (http://www.ief.es/recursos/estadisticas/fuentes_tributarias.aspx)

- No representatividad de no declarantes que no soportaron retención por rendimientos del trabajo.

La relevancia de estas limitaciones en nuestro estudio, sobre todo la relativa a la imposibilidad de construir declaraciones conjuntas a partir de individuales ni unir a los declarantes en hogares, debido a la inexistencia de información que las relacione, nos llevan a descartar esta fuente de información en el desarrollo de esta tesis doctoral.

3.4.2. El Panel de declarantes de IRPF

A finales de la década de los años ochenta el Instituto de Estudios Fiscales (IEF) abordó el primer Panel de Declarantes del Impuesto sobre la Renta de las Personas Físicas (IRPF), que inicialmente abarcó los ejercicios impositivos comprendidos entre 1982 y 1994, siendo posteriormente prolongado, en 2006, hasta el ejercicio de 1998. La experiencia muy positiva obtenida en la utilización del Panel de Declarantes de IRPF 1982-1998, junto con el auge internacional de los trabajos que utilizan bases de microdatos fiscales con estructura de panel, motivaron al IEF en el año 2005 a dar continuidad al proyecto iniciado hace veinte años mediante la difusión de microdatos de panel desde 1999 –año de entrada en vigor de la Ley 40/1998, que reformó con profundidad el IRPF español– hasta el año más reciente posible para el que se disponga de microdatos. Actualmente, está disponible el nuevo Panel de Declarantes de IRPF 1999-2012, construido a partir de la extracción de muestras de declaraciones de IRPF proporcionadas por la AT. Esta utilidad ha sido reconocida por el Consejo Superior de Estadística, habiéndose incluido como una Operación Estadística del Plan Estadístico Nacional, lo que supone dotar al proyecto de una continuidad en el tiempo, además de la obligación de realizar la difusión de acuerdo a los compromisos adquiridos.

Entre las características que definen el Panel de Declarantes de IRPF 1999-2012 podemos citar como más relevantes las siguientes:

- a) El año de arranque del Panel es el 1999, fecha clave por tratarse del primer ejercicio de implantación de la Ley 40/1998 que reformó sustancialmente la estructura del IRPF español vigente hasta 1998.
- b) El año 2003 es el año base del panel por tratarse del primer ejercicio de implantación de la Ley 46/2002 (que reformó parcialmente el IRPF vigente desde 1999) y el primer año en el que la AT grabó las declaraciones del impuesto hasta el nivel de identificación de las casillas de tres dígitos.
- c) La población objeto de estudio está constituida por el conjunto de declarantes del Impuesto sobre la Renta de las Personas Físicas en cada uno de los sucesivos ejercicios comprendidos entre 1999 y 2012.
- d) El ámbito geográfico de la investigación es el Territorio Fiscal Común, es decir, todo el territorio español con excepción de las Comunidades Autónomas forales del País Vasco y Navarra, las cuales tienen un sistema tributario propio gestionado por sus propias administraciones tributarias.

- e) Las unidades de análisis son las declaraciones de IRPF. De acuerdo con la legislación aplicable a los periodos que abarca el Panel, en el IRPF español se pueden diferenciar dos tipos de declaraciones según el número de declarantes: las realizadas en aplicación del régimen de tributación individual, en las que existe un único declarante, y las realizadas bajo el régimen de tributación conjunta, en las que bien son declarantes los dos cónyuges de los matrimonios que optan por éste régimen (conjuntas de matrimonios) o bien son declarantes la madre o el padre solteros, viudos o separados y sus hijos menores de edad o incapacitados judicialmente (conjuntas mono-parentales). En las conjuntas de matrimonios, el Panel ofrece información sobre los dos cónyuges, pero en las declaraciones conjuntas mono-parentales sólo ofrece información sobre el padre o la madre.
- f) El método de muestreo utilizado es un Muestreo estratificado aleatorio (tres niveles de estratificación) con afijación de mínima varianza (afijación de Neyman), con un error relativo de muestreo inferior al 1 por 1000 y un nivel de confianza adicional a del 1 por 1000. Los estratos son: la Comunidad Autónoma (quince de Territorio Fiscal Común más las Ciudades Autónomas de Ceuta y Melilla), la fuente de renta (si las rentas del trabajo son o no mayoritarias) y el nivel de renta (en una escala de diez tramos).
- g) El Panel de Declarantes de IRPF 1999-2012 se ha concebido como un panel expandido en el que cada nuevo ejercicio de IRPF se va incorporando al Panel teniendo en cuenta que los cortes transversales deben incluir, además de los declarantes presentes en ejercicios anteriores, una representación de las altas de declarantes que se produzcan en la población de ese año que compensen en términos de representatividad transversal las bajas que se van produciendo por desgaste del panel.
- h) Por último, hay que decir que, a diferencia de las muestras anuales de declarantes de IRPF, este Panel de Declarantes permite en el caso de matrimonios que optan por tributar de forma individual la vinculación de las dos declaraciones de los cónyuges. En relación con esta posibilidad, además de los factores de elevación de las variables a la población declarante, el Panel incorpora factores de elevación por **hogares fiscales** (unidades contribuyentes integradas por declarantes no casados en tributación individual, declarantes no casados con hijos menores en tributación conjunta —unidades mono-parentales—, matrimonios en tributación conjunta y matrimonios con cónyuges que tributan individualmente).

El Panel de IRPF 1999-2012 presenta una serie de potencialidades y limitaciones respecto de otras bases, bien las tradicionales basadas en encuestas, bien respecto de otras bases de microdatos de carácter fiscal alternativas, como son las Muestras Anuales de IRPF.

Entre las diversas potencialidades de utilizar la información contenida en los registros administrativos fiscales para trabajos orientados hacia la distribución y redistribución de la renta o en el desarrollo de micro-simuladores, destacan especialmente las asociadas al carácter censal de la población de referencia y el grado de detalle y precisión con el que suelen reflejarse las variables que

se incluyen en estas bases. No obstante, su dependencia de la normativa tributaria les confiere, algunas limitaciones, destacando las asociadas a la unidad observacional, distinta de la habitual definición de hogar económico usada en las encuestas, o las derivadas de la definición fiscal de las magnitudes.

Centrándonos en el estudio de la distribución y redistribución de la renta por ser el tema que nos ocupa en esta tesis, esta comparación puede realizarse respecto de encuestas como la Encuesta de Condiciones de Vida, la Encuesta de Presupuestos Familiares -ambas realizadas por el INE- o la Encuesta Financiera de las Familias, realizada por el Banco de España. Así, vemos, en primer lugar, que el origen informativo del Panel proporciona gran precisión a la hora de reflejar los distintos componentes de la renta personal, si bien es cierto, que bajo definiciones derivadas de la normativa fiscal. En el caso de las encuestas, el grado de detalle depende de cada una de ellas. Así, mientras las EPFs ofrecen una única cifra de renta del hogar, neta de impuestos y transferencias, la ECV detalla la composición de esa renta neta según se trata de rentas salariales, empresariales o del capital. La consideración en términos netos de los ingresos, es decir, después del pago de IRPF y de las cotizaciones sociales del empleado, introduce una limitación alternativa en las encuestas, pues exigen la elevación al importe íntegro de los mismos, con la dificultad que entraña dada la progresividad y ajustes de la carga fiscal del IRPF. Además, hay que tener en cuenta frente al carácter censal de los registros administrativos, que la infra-representación y la falta de respuesta son problemas habituales presentes en las encuestas. En cuanto a la riqueza informativa de sus datos, las bases de microdatos fiscales permiten realizar muchos análisis que están vedados a otras muestras de origen no fiscal, especialmente en el ámbito impositivo, dada la precisión con la que se definen los distintos componentes monetarios de la estructura del impuesto. Este hecho resulta crucial para la construcción de modelos de microsimulación impositiva, donde la liquidación de impuestos de estructura compleja como el IRPF, a partir de la información contenida en bases de microdatos como la ECV, la EPF o la EFF, introducen sesgos importantes en las proyecciones generadas.

Pero sin duda, la principal limitación de las bases de microdatos fiscales, que está presente tanto en el Panel como en las Muestras Anuales de Declarantes de IRPF, y que nos llevan a no proseguir en el estudio de las mismas a lo largo de este trabajo de investigación, es la unidad observacional considerada, distinta del concepto tradicional de hogar económico empleado por las encuestas. Aunque sabemos que estas bases fiscales cuentan con información sobre descendientes y ascendientes, éstos siempre están vinculados a una unidad contribuyente, bien sea un matrimonio o un individuo no casado. En cambio, no es factible vincular distintos contribuyentes adultos cuando éstos no están casados, aunque exista convivencia. Sin embargo, esta limitación no tiene por qué ser invalidante, como puede verse en muchos de los trabajos que utilizan este tipo de bases de microdatos fiscales, tanto referidos a la realidad española como a la de otros países con bases disponibles similares. En unos casos, la definición de un concepto de hogar fiscal en el que agregar las rentas y cuotas impositivas de los cónyuges resulta suficiente para evaluar las distribuciones correspondientes, incluso pudiéndose aplicar al mismo escalas de equivalencia referidas a los adultos e

hijos menores presentes en cada unidad familiar contribuyente. En otros casos, en los que resulta indispensable considerar el hogar económico como unidad de análisis, la utilización complementaria del Panel junto a encuestas como la ECV o la EPF, consideradas como bases receptoras en procesos de fusión estadística (*statistical matching*), cada vez más extendidos, permiten salvar esta limitación, sin renunciar a la riqueza informativa de los microdatos fiscales. En cualquier caso, defendemos aquí la idea cada vez más extendida de la complementariedad entre ambos tipos de bases de datos.

3.5. La Muestra Continua de Vidas Laborales

La Muestra Continua de Vidas Laborales (en adelante MCVL) constituye una fuente de información inestimable que desde el año 2004, por iniciativa de la Secretaría de Estado de la Seguridad Social del extinto Ministerio de Trabajo e Inmigración (MTIN), nos ofrece datos relativos de la distribución de la población de un determinado año según diversas características laborales registradas administrativamente a través de la Seguridad Social, así como de su historia laboral desde que existen registros informatizados.

En concreto, la MCVL es un extracto de datos individuales anonimizados, procedentes de las bases de datos de la Seguridad Social, a los que se añaden otros -mediante procesos de fusión estadística- que se toman del Padrón Municipal Continuo (INE) y, dependiendo de las versiones, del resumen anual de retenciones e ingresos a cuenta del IRPF (Modelo 190) de la AT. Estos microdatos están referidos a algo más de un millón de personas y constituyen una muestra representativa de todas las personas que tuvieron relación con la Seguridad Social en un determinado año. Se denomina continua porque se actualiza anualmente, y es de vidas laborales porque cada muestra, aunque referida a la población cotizante o pensionista en el año de referencia, reproduce el historial anterior de las personas seleccionadas para las variables procedentes de la Seguridad Social, remontándose hacia atrás hasta donde se conserven registros informatizados. Es decir, cada año permanecen en la muestra los que ya estaban y han seguido teniendo relación con la Seguridad Social, y se incorporan a la muestra nuevas personas que por primera vez son cotizantes o pensionistas, con los mismos criterios de selección. Por ello, las sucesivas versiones de la muestra funcionan como un panel, cubriendo en la actualidad el periodo que va desde la primera extracción aleatoria del año 2004 hasta que el año más reciente completo, 2015.

Los objetivos que han llevado a la creación de la MCVL son el apoyo a la investigación y la transparencia. La Seguridad Social requiere para su gestión registrar información sobre la vida laboral de los distintos individuos que o bien reciben prestaciones, o bien contribuyen con sus cotizaciones. Esta información individualizada se conserva durante muchas décadas, lo que abre la posibilidad de desarrollar estudios que anteriormente sólo podían ser abordados con información parcial o de carácter agregado. El tamaño de la muestra permite estudiar colectivos relativamente pequeños, que en otro tipo de encuestas no suelen ser significativos, al tiempo que los datos referidos al pasado de una misma persona facilitan los estudios dinámicos.

Cada edición anual de la MCVL toma como año de referencia el último año natural finalizado. A efectos de su inclusión en una edición anual de la MCVL, la persona debe disponer de un documento Identificador de Persona Física¹⁹ (IPF) cuyo número pertenezca al conjunto de números seleccionables - el cual incluye el 4% de todos los números posibles-; y haber formado parte de la población de referencia, como afiliada o como pensionista, al menos un día dentro del año de referencia. La población de referencia está formada por las personas en situación de afiliado en alta, ya sea laboral, en convenio especial, o como perceptor de prestación o subsidio por desempleo o cese de actividad, y las beneficiarias de alguna pensión contributiva de la Seguridad Social. No se incluyen en ella quienes exclusivamente están registrados en la Seguridad Social a efectos de recibir asistencia sanitaria, ni los perceptores de pensiones no contributivas o asistenciales, ni quienes están encuadrados en sistemas de previsión social distintos de la Seguridad Social, como es el caso de los funcionarios de Clases Pasivas, ni los inscritos como demandantes de empleo que no perciben prestación o subsidio por desempleo, ni tampoco los perceptores de la Renta Activa de Inserción.

Por lo que se refiere al ámbito temporal, la población de la MCVL, definida como todos los individuos que han tenido relación con la Seguridad Social en algún momento del año, es de mayor tamaño y tiene una composición algo diferente respecto a lo que se obtendría a una fecha fija, que es el criterio habitual. La elección de una fecha fija limita la presencia de personas que entran y salen de la situación de alta laboral sin permanecer mucho tiempo en ella, incluso cuanto se toman medias, pero especialmente si dicha fecha es el final del año. Es un grupo bastante numeroso: se estima que alrededor de tres millones de personas, habiendo trabajado en algún momento del último año, no están de alta en un día determinado. Estas personas, que son mayoritariamente mujeres y jóvenes y tiene contratos de corta duración, no hubieran formado parte de la población si ésta se hubiera definido en referencia a un día determinado. Por tanto, los correspondientes valores de edad, sexo o duración de contrato estarán mucho más representados en la MCVL que en otras fuentes de información, como son los datos publicados de afiliación o la Encuesta de Población Activa²⁰.

Uno de los aspectos que ofrece al mismo tiempo mayores oportunidades, es que cada persona puede pertenecer a varias categorías distintas, simultánea o sucesivamente. Las personas para las que se ofrecen datos no pueden ser divididas simplemente en pensionistas y cotizantes: algunos eran las dos cosas al mismo tiempo. Tampoco se puede dividir la población de manera nítida por zonas geográficas: una misma persona puede aparecer como residente en un sitio, pero trabajar en otro, o haberlo hecho en el pasado.

Se realizan dos versiones anuales de la MCVL, una sin datos fiscales (SDF) y otra con ellos (CDF) que aparece unos meses más tarde.

La MCVL SDF contiene una serie de variables procedentes de la Seguridad Social, de las que algunas son simplemente identificativas (de la persona, del código de cuenta de cotización, de la empresa, del año a que se refiere el

¹⁹ Generalmente un Documento Nacional de Identidad (DNI) o N° de Identificación de Extranjero (NIE). En ocasiones, como en el caso de pensionistas extranjeros o pensionistas menores de edad, puede ser otro tipo de número.

²⁰ Una comparación entre la EPA (Encuesta de Población Activa) y la MCVL puede verse en Ramos (2007).

dato...). Otras se refieren a la persona: fecha de nacimiento y fallecimiento si lo hubiera, sexo, nacionalidad, provincia donde se afilió por primera vez. Doce variables tienen que ver con el trabajo desempeñado: régimen, fechas de alta o baja, tipo de contrato, tipo de jornada, grupo de cotización, causa de baja, minusvalía si consta en el contrato, información similar para otros regímenes y situaciones de afiliación, y base por la que se ha cotizado cada mes en cada relación laboral, que es la variable que más espacio ocupa. Otras cinco variables recogen modificaciones que puedan haberse producido durante la relación laboral. Diez variables se refieren al empleador o responsable de la cotización: actividad económica, tamaño, antigüedad como empleador, dos referidas a su localización y cinco que ayudan a describir el tipo de empleador. Para los que ya sean pensionistas, una treintena de variables describen la pensión: clase, grado y fecha de revisión de minusvalía para las de incapacidad, régimen, fecha de inicio y cese, base reguladora, porcentaje aplicado y pensión inicialmente calculada, revalorización acumulada, complemento al mínimo, y otros importes, pagas extras, pensión total mensual y de cada año, causa de cese si lo hubiera, y una serie de características relevantes solo para algunos tipos de pensiones.

El Padrón Municipal Continuo aporta información más precisa que la que se dispone en los registros de Seguridad Social sobre el lugar de nacimiento (país o provincia) y el lugar de residencia. Se ha incorporado también a la Muestra el nivel educativo, pero, como se indica en la ficha correspondiente, esta variable debe utilizarse con muchas precauciones porque no se actualiza debidamente. Además, el Padrón aporta información sobre la edad y el sexo de las personas inscritas en la misma hoja padronal que la persona seleccionada para la Muestra, hasta un máximo de diez.

Las versiones de la muestra con datos fiscales –MCVL CDF- incluyen, además de variables procedentes de Seguridad Social y del Padrón, las contenidas en el informe que, al finalizar cada año, los empleadores remiten a la Agencia Tributaria sobre las Retenciones del Impuesto sobre la Renta de las Personas Físicas (IRPF) derivadas de trabajo por cuenta ajena para dicho ejercicio (Modelo 190). Esto incluye 43 variables entre las que se encuentra la percepción íntegra, (es decir, el salario sin los topes que afectan a las bases de cotización de la Seguridad Social) tanto en dinero como en especie. También está la cuantía de las prestaciones públicas o privadas. Por otro lado, contiene información sobre la persona o la familia, como el estado civil, la discapacidad, las pensiones compensatorias pagadas al cónyuge, o el número de ascendientes y descendientes que dependen económicamente del interesado, con información de su edad y discapacidad. Por tanto, figuran en el fichero MCVL CDF las personas que forman parte de la Muestra y de las que se dispone de información sobre dichas retenciones. Si una persona de la Muestra no aparece en este fichero puede deberse a que no ha sido correctamente identificada, o a que no estaban en el campo de aplicación de estas Retenciones (Navarra y País Vasco; trabajo por cuenta propia; empleados de hogar, etc). Tampoco aparecen los trabajadores no incluidos en la Seguridad Social (como los funcionarios cuyo sistema de protección es el de Clases Pasivas y Mutualismo) aunque tengan retenciones para el IRPF.

A pesar de su riqueza económica, la MCVL adolece a la vez importantes limitaciones. Las más importantes están relacionadas con aquellas que presenta la información fiscal de la que se nutre la encuesta. Así, no forman

parte de la población objeto de estudio las personas para las que no se está obligado a rellenar el modelo 190: los residentes en Navarra y País Vasco, los trabajadores por cuenta propia, los empleados del hogar y los trabajadores que no cotizan a la Seguridad Social. Pero tampoco aparecen los trabajadores no incluidos en la Seguridad Social (como los funcionarios cuyo sistema de protección es el de Clases Pasivas y Mutualismo).

Otra de las limitaciones que nos impone la MCVL es la relativa a la información sobre las prestaciones ya que tan sólo se recogen las pensiones contributivas, pero en cambio no hay datos sobre pensiones no contributivas, ni sobre otras prestaciones de la Seguridad Social, como Maternidad o Incapacidad Temporal, durante las cuales el trabajador sigue encuadrado en su empresa y aparece como una relación laboral normal.

La relevancia que, por motivos obvios, tienen estas limitaciones en nuestro estudio, unidas a la unidad observacional considerada en esta fuente, tan distinta del concepto tradicional de hogar económico empleado por las encuestas y que pretendemos emplear en nuestro trabajo, hace que, a la hora de estimar una distribución de la renta combinada no vayamos a considerar la MCVL como una fuente de información que nos permita mejorar la información relativa al conjunto de todas las unidades pertenecientes al sector hogares.

4. Disponibilidad de los datos

Una vez analizada toda la información disponible sobre la renta de los hogares, tanto a nivel agregado como a nivel de microdatos, podemos justificar la selección de las fuentes con las que vamos a trabajar, así como el período de tiempo que va a resultar más óptimo para nuestro estudio.

Conscientes de que el marco de referencia de los datos sobre ingresos de las familias españolas es la Contabilidad Nacional de España (CNE), que como sabemos nos proporciona dicha información económica a nivel agregado, tenemos a nuestro alcance la serie de Cuentas no financieras del total de la economía y de los sectores institucionales en Base 2010, publicada por el INE, para el periodo 1999-2015²¹. De este conjunto de información podremos extraer para el sector hogares cada una de las partidas contables que intervienen en las distintas fases del proceso económico: producción, generación de la renta, distribución de la renta, redistribución de la renta, utilización de la renta y acumulación financiera y no financiera. En concreto, como ya sabemos, en este trabajo centraremos nuestra atención muy especialmente en aquellas partidas vinculadas al proceso de distribución de la renta de los hogares, cuya agregación da como resultado la Renta Disponible correspondiente a dicho sector, que será tenida en cuenta en todo momento en el proceso de estimación de la distribución personal de la renta a partir de datos individuales que llevaremos a cabo en este trabajo.

²¹ <http://www.ine.es/jaxi/menu.do?type=pcaxis&path=%2Ft35%2Fp008&file=inebase&L=0>

En cuanto a la disponibilidad de información económica de los hogares a nivel de microdatos, tendremos que:

- Disponemos de un período relativamente largo, una década, de una fuente de información homogénea que desagrega los distintos tipos de renta para cada hogar con un nivel de detalle apropiado para el estudio que abordamos en este trabajo, la ECV. Considerada como el principal instrumento diseñado por Europa para la medición de los estándares de vida de sus ciudadanos de forma comparable entre países, la ECV será nuestra fuente de microdatos base sobre la que iremos construyendo una distribución personal de la renta más acorde con los datos de CNE que la que se deriva de la propia encuesta. La última actualización de microdatos de la ECV base 2013 es de mayo de 2016²² y proporciona información de la ECV de 2015 (rentas de 2014).
- Las cifras de la EPF base 2006, que como veremos a lo largo de este estudio, nos aportarán información adicional sobre el nivel de ingresos de un hogar en base a la totalidad de sus gastos declarados, cubren actualmente el periodo 2006-2015²³, coincidiendo el año de realización de la encuesta con el periodo de referencia de los datos que recoge.
- La EFF²⁴, que se viene realizando desde el año 2002 con el fin de ofrecernos información directa sobre las condiciones financieras de las familias españolas, y que en este trabajo nos permitirá relacionar las rentas, los activos, las deudas y el gasto de cada una de ellas, se encuentra disponible también para los años 2005, 2008, 2011 y 2014²⁵. El período de referencia de los datos en este caso es el mismo que el año de la entrevista y, adicionalmente, se publica el conjunto de rentas no relacionadas con la actividad laboral percibidas en el año natural anterior (2010 para el caso de la EFF 2011).

En el cuadro 1.5 se presenta una visión de conjunto de la información disponible desde el año 2004 hasta la actualidad sobre las fuentes seleccionadas para su estudio a lo largo de esta investigación. Se puede observar cómo la coincidencia en el tiempo de todas las fuentes se encuentra condicionada por la frecuencia de ejecución de la EFF, siendo los años 2008 y 2011 en los que se tiene la realización de todas ellas. Así pues, en ambos años podremos contar tanto con la información de las antiguas encuestas de la ECV (base 2004) en las que la recogida de los ingresos se realizaba mediante cuestionarios, como de la nueva información que el INE ha reelaborado para estos mismos años según la nueva metodología (base 2013) empleando los pertinentes registros administrativos; además, estarán disponibles la EPF y la EFF que aportarán en dichos años información económica adicional respecto a las ECV. La selección de los años 2008 y 2011 nos va a permitir realizar no sólo un estudio de corte transversal en cada uno de ellos, sino también valorar los cambios ocurridos en esos tres años.

²²http://www.ine.es/dyns/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176807&menu=resultados&secc=1254736195153&idp=1254735976608

²³http://www.ine.es/dyns/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176806&menu=resultados&secc=1254736195147&idp=1254735976608

²⁴http://www.bde.es/bde/es/areas/estadis/Otras_estadistic/Encuesta_Financi/

²⁵ Los microdatos del año 2014 ha sido publicados por el Banco de España escasamente un mes antes de terminar la memoria de esta Tesis, por lo que ha sido materialmente imposible su consideración en este trabajo.

Cuadro 1.5
Disponibilidad de fuentes estadísticas desde 2004

	Año de referencia de la información											
	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
CNE	CNE ₁₀ ⁰⁴	CNE ₁₀ ⁰⁵	CNE ₁₀ ⁰⁶	CNE ₁₀ ⁰⁷	CNE ₁₀ ⁰⁸	CNE ₁₀ ⁰⁹	CNE ₁₀ ¹⁰	CNE ₁₀ ¹¹	CNE ₁₀ ¹²	CNE ₁₀ ¹³	CNE ₁₀ ¹⁴	CNE ₁₀ ¹⁵
ECV	ECV ₀₄ ⁰⁵	ECV ₀₄ ⁰⁶	ECV ₀₄ ⁰⁷	ECV ₀₄ ⁰⁸	ECV ₀₄ ⁰⁹	ECV ₀₄ ¹⁰	ECV ₀₄ ¹¹	ECV ₀₄ ¹²				
				ECV ₁₃ ⁰⁸	ECV ₁₃ ⁰⁹	ECV ₁₃ ¹⁰	ECV ₁₃ ¹¹	ECV ₁₃ ¹²	ECV ₁₃ ¹³	ECV ₁₃ ¹⁴	ECV ₁₃ ¹⁵	
EPF	ECPF ₇₇ ⁰⁴	ECPF ₇₇ ⁰⁶	EPF ₀₆ ⁰⁶	EPF ₀₆ ⁰⁷	EPF ₀₆ ⁰⁸	EPF ₀₆ ⁰⁹	EPF ₀₆ ¹⁰	EPF ₀₆ ¹¹	EPF ₀₆ ¹²	EPF ₀₆ ¹³	EPF ₀₆ ¹⁴	EPF ₀₆ ¹⁵
EFF		EFF ⁰⁵			EFF ⁰⁸			EFF ¹¹			EFF ¹⁴	
donde, CNE: Contabilidad Nacional de España. ECV: Encuesta de Condiciones de Vida. EPF: Encuesta de Presupuestos Familiares. (ECPF: Encuesta Continua de Presupuestos Familiares). EFF: Encuesta Financiera de las Familias.												

Fuente: Elaboración propia, a partir de la información disponible a 24 de enero de 2017

Una vez realizada una breve descripción sobre las principales fuentes de información de ingresos de los hogares que tenemos a nuestro alcance, y seleccionadas, de entre aquéllas, las que serán de mayor utilidad a lo largo de este estudio, en el siguiente capítulo profundizaremos en el estudio de la congruencia entre los ingresos de los hogares que se desprenden de la ECV, que como ya sabemos nos proporciona un amplio grado de especialización y detalle de datos de renta a nivel individual de hogar, y los procedentes de la CNE, el marco de referencia para los ingresos de los hogares medidos de forma agregada.

Capítulo II

**Los ingresos de los hogares: Congruencia
entre la Encuesta de Condiciones de Vida
y la Contabilidad Nacional de España**

1. Introducción

En los últimos años, muchos trabajos de índole socioeconómica han venido empleando los microdatos de las encuestas a hogares (*Encuestas de Presupuestos Familiares*, *Panel de Hogares de la Unión Europea*, *Encuestas de Condiciones de Vida*) que realiza el Instituto Nacional de Estadística de España (INE), con el fin de analizar las más diversas características de las distribuciones de las rentas de hogares y personas en nuestro país, prestando especial interés a sus niveles de desigualdad y pobreza, entre otros aspectos.

El INE procura la máxima calidad de sus microdatos, vigilando desde su diseño hasta el último de sus procesos de producción; y la confianza de los usuarios en su tradicional buen hacer explica que se hayan venido utilizando sistemáticamente, asumiendo para ello la hipótesis de que representan fielmente las rentas de los hogares españoles. Sin embargo, la natural reticencia humana (y, por lo tanto, de muchos encuestados) a comunicar a los demás sus datos personales de tipo económico, puede conducir a una recogida de información no siempre fiel a la realidad de las personas y de sus hogares. Y posiblemente, como consecuencia de ello, diversos autores (Sanz, 1995; Pena *et al.*, 1996; Delgado y Mercader, 2001) han venido constatando ciertas divergencias entre la información que proporcionan estos microdatos y la que sobre las rentas de los hogares españoles presenta la Contabilidad Nacional de España -CNE- no solo a nivel agregado, sino también en relación con los diferentes tipos de rentas consideradas.

Por otra parte, es evidente que la naturaleza de los ingresos de los hogares no es idéntica en toda la población, sino que se asocia fuertemente para diferentes grupos poblacionales con ciertas características comunes. Así, se precisa ser asalariado en algún momento para percibir rentas salariales del trabajo; o ser empresario o profesional libre para recibir rentas del trabajo por cuenta propia; o estar parado para percibir prestaciones sociales por desempleo... Y, por lo tanto, niveles de ocultación en las declaraciones de determinados tipos de ingresos que pudieran ser muy distintos –o desequilibrados– para los diversos grupos poblacionales podrían producir sesgos presumiblemente importantes en las mediciones de fenómenos sociales, como la desigualdad o la pobreza, mediante indicadores métricos objetivos.

De este modo, y en caso extremo, si en una cierta subpoblación se ocultaran sistemáticamente la totalidad de sus ingresos reales, declarando ingresos nulos, todos los pertenecientes a esa subpoblación se situarían con casi total seguridad bajo los umbrales de pobreza calculados para la población en su conjunto por métodos objetivos. Sin embargo, los integrantes de esa subpoblación no tienen por qué ser necesariamente pobres, pues simplemente han ocultado sus rentas. Pero no es este el único efecto que se produciría, ya que la ocultación de rentas de esta subpoblación sesgaría a la baja los indicadores de posición central de la distribución de ingresos (medias, mediana) y, en consecuencia, sesgaría también a la baja el umbral de la pobreza, expulsando de la categoría de pobres a otros perceptores de rentas bajas, verdaderos pobres si declararon sus rentas reales.

En consecuencia, aun cuando se calculen cuidadosamente este tipo de indicadores objetivos universalmente aceptados, ciertas afirmaciones que se suelen realizar de forma contundente (como es el caso de “el umbral de la pobreza de tal población es tanto”, o “tal población es más desigual que tal otra”, o “tal población tiene una tasa de pobreza mayor que tal otra”) posiblemente no deberían hacerse sin investigar previamente cuáles son las poblaciones a las que realmente se están refiriendo, ni sin analizar antes si las fórmulas aplicadas sobre los microdatos miden adecuadamente los niveles del fenómeno estudiado.

Así pues, el primer objetivo que nos planteamos en este capítulo es evaluar la congruencia –o el grado de acuerdo existente– entre los datos de renta de los hogares proporcionados para los años²⁶ 2008 y 2011 por las Encuestas de Condiciones de Vida (ECV) del INE –realizadas en 2009 y 2012, respectivamente–, y la información correspondiente facilitada por la CNE (base 2010) relativa a dicho período. Para ello, en el apartado 2 examinaremos la renta disponible proporcionada por la ECV tratando de encontrar rasgos que nos orienten en el análisis de las discrepancias observadas con la CNE a nivel agregado. En la sección 3, intentando profundizar en ese análisis mediante el estudio del marco teórico de referencia –SEC 2010–, diseñaremos una Cuenta de Rentas del Sector Hogares que, con cierto nivel de desagregación, nos permitirá comparar de forma homogénea los datos de la ECV con los de la CNE, calculándola para ambas fuentes y comparando sus resultados (apartado 4).

2. La distribución de los tamaños de los ingresos en España según la ECV

Como ya anticipamos en el Capítulo 1, uno de los objetivos fundamentales de la ECV consiste en proporcionar información sobre la distribución del nivel de ingresos en España, y de sus componentes, de las personas y los hogares.

En concreto, la renta disponible total de los hogares viene recogida en dicha encuesta a través de la variable *vhRentaAla*²⁷. A partir de ella, se puede obtener la correspondiente renta per cápita en base 2004 para los años 2008 y 2011 junto a su equivalente²⁸ en CNE, basada en la Renta Disponible Neta (rúbrica B.6n) del sector hogares (cuadro 2.1). Observando este cuadro podemos detectar unas primeras discrepancias entre ambas fuentes. Así, la renta disponible per cápita que se deduce de la ECV, asciende a 11.103 euros en 2008 y 10.860 euros en 2011, y muestra importantes diferencias (infraestimación del 18,8% y 20,6%, en 2008 y 2011, respectivamente) con

²⁶ A pesar de las ventajas que nos ofrece la posibilidad de trabajar con datos de panel, la elección de la ECV como fuente que nos proporciona mayor información sobre ingresos, nos condiciona a trabajar con datos de corte transversal.

²⁷ *vhRentaAla* incluye el alquiler imputado, así como las rentas percibidas de esquemas privados de pensiones.

²⁸ Como veremos más adelante, la ECV no incluye el consumo de capital fijo en sus estimaciones, por lo que a lo largo de este trabajo las comparaciones de los saldos y agregados contables de CNE siempre se efectuarán en términos netos.

respecto a la que se deriva de la Cuentas de Renta de los hogares de la CNE (13.674 euros para 2008 y 13.673 euros para 2011).

Cuadro 2.1

RENTA DISPONIBLE NETA AÑOS 2008 Y 2011: CNE vs. ECV BASE 2004

	CNE 2008 - <i>B.6n</i>	ECV 2009 - <i>vhrentaAla</i>
Renta disponible neta de los hogares para España en el año 2008 (miles de euros)	628.767.000	509.312.525
Renta disponible neta per cápita de los hogares para España en el año 2008 (euros)	13.674	11.103
Población (nº de personas)*	45.983.168	45.870.296

	CNE 2011 - <i>B.6n</i>	ECV 2012 - <i>vhrentaAla</i>
Renta disponible neta de los hogares para España en el año 2011 (miles de euros)	639.038.000	503.398.599
Renta disponible neta per cápita de los hogares para España en el año 2011 (euros)	13.673	10.860
Población (nº de personas)*	46.736.256	46.354.779

Fuente: Elaboración propia a partir de los datos del INE de la CNE Base 2010 para los años 2008 y 2011 y de las ECV Base 2004 relativas a los años 2009 y 2012, respectivamente.

Desde su implantación en 2004, la ECV había utilizado como método de recogida principalmente la entrevista personal. El propio INE, consciente de la subdeclaración de ingresos en la ECV, adoptó en 2013 una nueva metodología en la producción de los datos relativos a los ingresos del hogar basada en el uso de ficheros administrativos (base 2013), lo que ha revertido en una mejora de la calidad de los datos y en la eficiencia en el método de recogida, como podemos comprobar en el cuadro 2.2. Con esta nueva metodología, los porcentajes de infraestimación de la renta disponible per cápita de los hogares en la ECV respecto a CNE se reducen 12,3 puntos en 2008, hasta el 6,5%, y 10,7 puntos en 2011, situándose en el 9,8%.

Cuadro 2.2

RENTA DISPONIBLE NETA AÑOS 2008 Y 2011: CNE vs. ECV BASE 2013

	CNE 2008 - <i>B.6n</i>	ECV 2009 - <i>vhrentaAla</i>
Renta disponible neta de los hogares para España en el año 2008 (miles de euros)	628.767.000	586.680.909
Renta disponible neta per cápita de los hogares para España en el año 2008 (euros)	13.674	12.790
Población (nº de personas)*	45.983.168	45.870.296

	CNE 2011 - <i>B.6n</i>	ECV 2012 - <i>vhrentaAla</i>
Renta disponible neta de los hogares para España en el año 2011 (miles de euros)	639.038.000	571.430.997
Renta disponible neta per cápita de los hogares para España en el año 2011 (euros)	13.673	12.327
Población (nº de personas)*	46.736.256	46.354.779

Fuente: Elaboración propia a partir de los datos del INE de la CNE Base 2010 para los años 2008 y 2011, de las Cifras de población a 1 de julio de 2015 y de las ECV Base 2013 relativas a los años 2009 y 2012, respectivamente.

(*) La CNE emplea como cifra de población la considerada a 1 de julio en las Estimaciones de la Población Actual de España, mientras que la ECV utiliza la derivada de las ponderaciones de los hogares multiplicados por el número de miembros de cada hogar.

De acuerdo con la metodología²⁹ de la ECV, la variable *vhrentaAla* se construye agregando la serie de componentes del ingreso que se muestran en el cuadro 2.3, los cuales se describirán en el siguiente apartado, así como los que constituyen la Renta Disponible Neta (*B.6n*).

Para profundizar en la comparación de la ECV y la CNE a nivel de componentes de renta, hemos construido los cuadros 2.4 y 2.5 (ECV base 2004 y ECV base 2013, respectivamente), en la que presentamos resumidamente los resultados comparativos obtenidos con los cálculos que hemos realizado para cada una de las componentes recogidas en el cuadro 2.3.

En el cuadro 2.4, observamos que, además de la sistemática infradeclaración en que incurre la ECV en los años estudiados, las Rentas de la propiedad, es decir, aquéllas que según la encuesta reciben los hogares a cambio de proporcionar fondos o poner sus activos materiales no producidos a disposición de otras unidades institucionales, presentan porcentajes de infradeclaración significativamente superiores al 66%, concretamente del 66,22% en 2008 y del 74,85% en 2011. Les seguirían muy de cerca las *Rentas de los trabajadores por cuenta propia*, con porcentajes de infradeclaración del 65,98% y 57,01% en 2008 y 2011, respectivamente. En otro bloque se encontrarían los componentes de la renta de la ECV que presentan porcentajes de infradeclaración entre el 40% y el 50% aproximadamente, como son las *Transferencias corrientes percibidas* (42,82% y 39,42%) y *pagadas* (47,98% y 50,55%). Por último, las rentas que mejor se recogen en la ECV para el periodo analizado son las *Rentas de los trabajadores por cuenta ajena*, con tasas de infradeclaración del 22,44% en 2008, y del 27,42% en 2011.

En sentido contrario, podemos ver que el *Alquiler imputado* se encuentra en principio sobreestimado respecto a CNE, un 38,92% en 2008, y un 18,07% en 2011. Por su parte, las *Otras rentas percibidas por los menores de dieciséis años* y el componente *Pago de intereses* no se encuentran con este desglose en la CNE, motivo por el cual no podemos profundizar en su análisis.

²⁹ INE (2005): *Encuesta de Condiciones de Vida: Metodología*. Madrid, INE.

Cuadro 2.3
COMPONENTES DE LA RENTA DISPONIBLE DEL HOGAR EN LA ECV

Variable	Descripción
	Renta bruta del asalariado
PY010G	(+) Renta bruta monetaria o cuasi monetaria del asalariado
PY020G	(+) Renta bruta no monetaria del asalariado
PY030G	(+) Cotizaciones sociales brutas a cargo del empleador
	Renta del trabajador por cuenta propia
PY050G	(+) Beneficios o pérdidas monetarios brutos de trabajadores por cuenta propia (incluidos derechos de propiedad intelectual)
PY070G* ó HY170N**	(+) Valor de los bienes producidos para autoconsumo
	Alquiler imputado
HY030N	(+) Alquiler imputado (sólo para vivienda principal)
	Rentas de la propiedad
HY040G	(+) Renta bruta procedente del Alquiler de una propiedad o terreno
HY090G	(+) Intereses, dividendos y ganancias de inversiones de capital en empresas no constituidas en sociedad (brutos)
	Transferencias corrientes percibidas
	<i>PRESTACIONES SOCIALES</i>
PY090G	(+) Prestaciones por desempleo (brutas)
PY100G	(+) Prestaciones por jubilación (brutas)
PY110G	(+) Prestaciones por supervivencia (brutas)
PY120G	(+) Prestaciones por enfermedad (brutas)
PY130G	(+) Prestaciones por invalidez (brutas)
PY140G	(+) Ayudas para estudios (brutas)
HY050G	(+) Ayudas por familia/hijos (brutas)
HY060G	(+) Exclusión social no clasificada en otro apartado (bruta)
HY070G	(+) Ayudas para vivienda (brutas)
	<i>TRANSFERENCIAS PERIÓDICAS PERCIBIDAS DE OTROS HOGARES</i>
HY080G	(+) Transferencias periódicas monetarias brutas percibidas de otros hogares
	Transferencias corrientes pagadas
PY030G	(-) Cotizaciones sociales a cargo del empleador
HY140G	(-) Impuesto sobre la renta y cotizaciones sociales (bruto)
HY120G	(-) Impuesto sobre el patrimonio
HY130G	(-) Transferencias periódicas monetarias brutas abonadas a otros hogares
	Otras rentas percibidas
HY110G	(+) Renta bruta percibida por los menores de 16 años
	Pago de intereses
HY100G	(-) Intereses pagados de préstamos hipotecarios (brutos)

(*) Hasta la ECV 2009

(**) A partir de la ECV 2010

Fuente: Elaboración propia a partir de la ECV

Cuadro 2.4
**ERRORES RELATIVOS DE LOS COMPONENTES DE INGRESO
 DE LA ECV (BASE 2004) SOBRE CNE**

	2008	2011
Renta bruta del asalariado	-22,44%	-27,42%
Renta del trabajador por cuenta propia	-65,98%	-57,01%
Alquiler imputado	38,92%	18,07%
Rentas de la Propiedad	-66,22%	-74,85%
Transferencias corrientes percibidas	-42,82%	-39,42%
Otras rentas percibidas (<16 años) ¹	-	-
Pago de intereses ²	-	-
Transferencias corrientes pagadas	-47,98%	-50,55%

(1) En CNE no existe una partida específica para recoger este tipo de rentas.

(2) Como veremos más adelante, en CNE esta variable se incluye en el Excedente Bruto de Explotación del Alquiler imputado como Consumos Intermedios.

Fuente: Elaboración propia a partir de los datos del INE de la CNE Base 2010 y los microdatos de las ECV (base 2004).

En el cuadro 2.5, el mismo ejercicio de comparación realizado con la ECV base 2013 ofrece una mejora en los niveles de sub-declaración en todos los componentes de renta en esta situación en base 2004 para el periodo de estudio. Así, por ejemplo, la incorporación de los registros administrativos para la producción de los datos de renta de la ECV ha reducido en el año 2008 el porcentaje de infradeclaración en más de 50 puntos porcentuales (pp) en las *Rentas de propiedad* y en casi 27pp las *Rentas del trabajador por cuenta propia* en el mismo periodo, hasta el 15,26% y 39,39%, respectivamente. Las *Transferencias corrientes pagadas* han pasado de registrar un porcentaje de infradeclaración del 42,82% en 2008 a situarlo en 35,53% en dicho año. Y las *Transferencias corrientes percibidas* y las *Rentas por cuenta ajena* han recortado alrededor de 9pp de distancia con respecto a CNE, alcanzando en la base 2013 un porcentaje de infradeclaración del 33,82% las primeras, y del 12,81% las segundas en 2008. En 2011, la mejora obtenida ha sido similar en cada uno de los componentes de la renta de la ECV analizados.

Cuadro 2.5
**ERRORES RELATIVOS DE LOS COMPONENTES DE INGRESO
 DE LA ECV (BASE 2013) SOBRE CNE**

	2008	2011
Renta bruta del asalariado	-12,81%	-15,05%
Renta del trabajador por cuenta propia	-39,39%	-39,16%
Alquiler imputado	38,92%	18,07%
Rentas de la Propiedad	-15,26%	-32,83%
Transferencias corrientes percibidas	-33,82%	-30,29%
Otras rentas percibidas (<16 años) ¹	-	-
Pago de intereses ²	-	-
Transferencias corrientes pagadas	-35,53%	-34,97%

(1) En CNE no existe una partida específica para recoger este tipo de rentas.

(2) Como veremos más adelante, en CNE esta variable se incluye en el Excedente Bruto de Explotación del Alquiler imputado como Consumos Intermedios.

Fuente: Elaboración propia a partir de los datos del INE de la CNE Base 2010 y los microdatos de las ECV (base 2004).

Las discrepancias observadas entre ECV y CNE no se encuentran sólo a nivel nacional, ya que si profundizamos en la comparación entre dichas fuentes a nivel de comunidades autónomas (CC.AA.) y las ordenamos decrecientemente en función de sus rentas medias per cápita recogidas en la Contabilidad Regional de España, como hacemos en (Callealta y Romero, 2013a), podemos observar cómo:

“además de la sistemática infradeclaración en que incurre la ECV, cambian sustancialmente su lugar en la ordenación comunidades como País Vasco, Navarra, La Rioja, Asturias, Ceuta o Murcia, entre otras. Sin embargo, es difícil valorar si existen o no anomalías en los indicadores relativos a la desigualdad y la pobreza, ya que las referencias al respecto en la literatura han sido generadas habitualmente con estos datos” (Callealta y Romero, 2013a).

Asimismo, si seguimos analizando los datos suministrados por la ECV para distintas clasificaciones de las personas encuestadas basadas en su sexo, nivel de estudios, grupos de edad, nacionalidad y tipo de hábitat, no encontramos evidencias que contraríen singularmente la intuición generalizada sobre sus comportamientos. En concreto, de aquel trabajo³⁰, se desprende que:

- *“Las rentas per cápita de los hogares donde habitan los hombres (9.769€ per cápita) son ligeramente superiores a las de los hogares donde habitan las mujeres (9.635€ per cápita), con similares tasas de desigualdad y pobreza”.*
- *“Las rentas per cápita crecen claramente con el nivel de estudios de las personas (6.959€ para “Sin estudios”, 8.219€ para “Educación Primaria”, 8.429€ para “Educación Secundaria 1ª etapa”, 10.294€ para “Educación Secundaria 2ª etapa” y 13.923€ para “Estudios superiores”), a la vez que disminuyen las correspondientes tasas de pobreza (39,1%, 27,0%, 23,9%, 17,0% y 8,7%, respectivamente), manteniéndose la desigualdad en niveles similares (en torno al 30%) en todas las categorías de personas con algún tipo de estudios, mientras que en la categoría “Sin estudios” se encuentra la mayor tasa de pobreza (25,2%)”.*
- *“Con la excepción del grupo de personas entre 25 y 35 años, las rentas per cápita muestran también la curva típica del ciclo vital de la renta, creciendo con la edad desde los 8.584€ en el grupo entre los 16 y 25 años, hasta los 11.240€ en el grupo de los 55-65, para decrecer hasta los 9.640€ para el grupo de más de 65 años. Sin embargo, el grupo de personas entre 25 y 35 años presenta una renta media más elevada de lo inicialmente esperado (11.124€), aunque no se trata de las rentas medias personales, sino de las rentas medias en los hogares de residencia de las personas y, probablemente, esta particularidad podría deberse a jóvenes que se pueden emancipar, quizás con una pareja que también percibe ingresos, y aún no tienen otras cargas familiares”.*
- *“Las rentas per cápita en los hogares donde viven los españoles (10.268€) son superiores a las de los hogares donde habitan las personas de los 24 países restantes de la Unión Europea (7.907€), y éstas a su vez son superiores a las de los hogares donde habitan las personas del resto*

³⁰ En Callealta y Romero (2013a) se estudió la congruencia entre los datos de la ECV 2010 base 2004 y la CNE 2009 base 2008.

del mundo (6.676€), con similares índices de desigualdad en todas las categorías y aumento paralelo de las tasas de pobreza (19,3%, 26,4% y 36,1%, respectivamente)".

- "Y las rentas per cápita de los que habitan en hogares localizados en zonas muy pobladas (10.772€) son superiores a las de los que habitan en zonas medianamente pobladas (9.053€), y éstas a su vez son superiores a las de los que viven en zonas poco pobladas (8.159€), con similares índices de desigualdad (en torno al 33%) y aumento paralelo de las tasas de pobreza (16,3%, 22,9% y 28,0%, respectivamente)".

Sin embargo, sí se encuentran en la ECV algunas evidencias singulares que parecen contrariar nuestra intuición. En general, en Callealta, F.J. y Romero, C. (2013a) hemos observado que:

- "Para las distintas categorías de empresarios, la de empresarios agrarios con asalariados presentan una renta media per cápita parecida a la de los inactivos, mientras que el resto de empresarios presentan rentas medias inferiores a las de los empleados no cualificados y a las de los parados".
- "Las cuatro categorías presentan participaciones en la renta total inferiores a su participación en la población española; y también mayores tasas de pobreza y de desigualdad que el resto de categorías".
- "Cuando se clasifican por la situación con relación a la actividad (variable PL031), los trabajadores por cuenta propia presentan rentas per cápita menores que los asalariados (incluso a tiempo parcial) y los jubilados".
- "Con relación a la situación profesional (variable PL040), los empleadores y empresarios sin trabajadores presentan menor renta per cápita que los asalariados (la mitad, en el segundo caso)".
- "Las rentas medias de los empleados por cuenta propia son sistemáticamente menores que las de los asalariados con la misma ocupación (variable PL050), tanto para los con dedicación a tiempo completo, como parcial".

Todos estos indicios parecen estar informándonos de una presumible infradeclaración de ingresos en la ECV en estos colectivos, de forma que podrían poner en cuestión los análisis realizados directamente sobre estos microdatos acerca de la distribución personal de las rentas en España, su grado de desigualdad y, pensamos que más críticamente aún, sobre la pobreza. Y para investigar si esta presumible infradeclaración de las citadas clases es, o no, la única causante de las discrepancias observadas a nivel global y autonómico, compararemos en los próximos apartados, desde la perspectiva del análisis insitucional, las distribuciones de rentas suministradas por ambas fuentes consideradas en este trabajo, la ECV y la CNE.

Asimismo, los errores en la medición de los componentes de la renta según la ECV en relación a la CNE, expuestos en los cuadros 2.4 y 2.5, nos llevan a plantearnos la necesidad de profundizar en las definiciones de los conceptos que subyacen en cada uno de ellos en una y otra fuente, con el objeto de construir con la encuesta una Cuenta de Rentas del Sector Hogares compuesta por partidas lo más parecidas posible a nivel teórico en ambas fuentes que, refinando la construida en nuestra anterior publicación (Callealta

y Romero, 2013a), nos permitirá llevar a cabo una comparación de los diferentes tipos de ingresos de los hogares lo más homogénea posible.

3. Marco teórico-conceptual para la comparación de ingresos de los hogares

A partir de las variables recogidas en la ECV (oleadas relativas a los años 2009 y 2012) y de la Cuentas de Rentas del Sector Hogares -CRSH- de la CNE (base 2010), para el año de referencia de los ingresos recogidos en aquéllas (2008 y 2011, respectivamente), el objetivo de este apartado es diseñar y construir, de forma razonada, una metodología que nos permita disponer de una CRSH para la ECV, por medio de la cual podamos calcular y comparar, con cierto nivel de desagregación y de forma lo más aproximada posible, sus resultados con los publicados por el marco contable de referencia (CNE). En consecuencia, dicha metodología, se empleará en analizar en profundidad la naturaleza de las discrepancias observadas en el apartado anterior entre la información proporcionada por ambas fuentes.

3.1. Análisis de la homogeneidad entre la ECV y la CNE

A partir de las variables recogidas en la ECV (oleadas relativas a los años 2009 y 2012) y de las Cuentas de Renta del Sector de Hogares de la CNE (base 2010), En este apartado llevaremos a cabo el análisis de las similitudes y diferencias entre los conceptos a emplear entre la CNE y la ECV desde la perspectiva de si ambas fuentes de información proporcionan datos comparables y que se encuentran disponibles para su estudio³¹, empezando por la Cuenta de asignación de la renta primaria de los hogares, hasta aquélla que nos permite obtener la Renta Disponible de dicho sector calculada en la encuesta, la Cuenta de distribución secundaria de la renta (véase cuadro 2.6). Así pues, estudiaremos en paralelo cómo define, por un lado, la CNE y, por otro, la ECV, cada una de las partidas que integran dichas cuentas.

Antes de ello, conviene hacer una aclaración muy importante en relación al significado de los conceptos de *bruto* y *neto* en que vienen expresadas las variables en cada una de las fuentes. Así, como ya mencionamos en el capítulo 1.2.1, en Contabilidad Nacional esta valoración afecta sólo a las magnitudes económicas que tienen la función de ser saldos contables, los cuales pueden expresarse en términos brutos o netos dependiendo de si incorporan o no, respectivamente, el Consumo de Capital Fijo (CCF). Por su parte, en la ECV las variables expresadas en términos brutos comprenden las cotizaciones sociales, cuando procedan, y las retenciones de impuestos que se les aplican, pero no incluyen el CCF en ningún caso ya que éste no se estima, salvo excepciones, en la encuesta. Asimismo, las variables “netas” de la encuesta llevarían descontadas las cotizaciones y retenciones de impuestos que procedan, así como el referido CCF. Dichas cotizaciones sociales en la

³¹ La CNE no publica los datos de estas cuentas con todo el nivel de detalle presentado en el capítulo I (cuadro 1.3) y, a su vez, la información que recoge la ECV tampoco es compatible con dicha desagregación.

ECV se refieren a las cotizaciones de asalariados (no se considera la cuota que corresponde a la empresa sino sólo a la del trabajador), trabajadores por cuenta propia y otros colectivos (en su caso) abonadas durante el período de referencia a los sistemas de seguros sociales obligatorios.

A su vez, entre las variables brutas de la ECV vamos a diferenciar, por un lado, entre aquellas que se corresponden con saldos contables de CNE: incluyen cotizaciones sociales y las retenciones de impuestos que procedan, pero llevan deducido el CCF y se corresponden a su vez con los saldos contables en términos netos de la CNE. Y, por otro, el resto de variables de la ECV en términos brutos que se corresponden con las variables de CNE que no funcionan como saldos contables entre las que, por tanto, la diferencia será únicamente las cotizaciones sociales y las retenciones de impuestos que se les apliquen en cada caso. Por tanto, de cara al análisis de la información a lo largo de este capítulo, y con el objeto de llevar a cabo una comparación entre los datos de las citadas fuentes lo más homogénea posible, a efectos de cálculos vamos a utilizar las variables de la ECV expresadas en términos brutos y las partidas contables netas siempre que proceda.

Una vez dicho esto podemos pasar ya a realizar la comparación de conceptos para cada una de las partidas que constituyen las cuentas que recogen nuestra información objeto de estudio: la **Cuenta de asignación de la renta primaria** y la **Cuenta de distribución secundaria de la renta**, relativas al sector hogares.

3.1.1. Análisis de la homogeneidad entre la ECV y CNE relativa a las variables de ingresos: Cuenta de asignación de la renta primaria.

En este apartado, estudiaremos cómo se definen, en el SEC 2010 y en la ECV, cada una de las partidas de la Cuenta de asignación de la renta primaria, es decir, *Excedente de explotación*, *Renta mixta*, *Remuneración de Asalariados* y *Rentas de la propiedad*.

EXCEDENTE DE EXPLOTACIÓN

En Contabilidad Nacional, el Excedente de Explotación Bruto (EBE) de una rama de actividad o sector institucional (i) es un saldo contable -de la Cuenta de explotación- que se corresponde con el Valor Añadido Bruto (VAB) después de pagar la Remuneración de Asalariados (RA) y los Otros Impuestos sobre la Producción netos de Subvenciones (OIN), es decir:

$$EBE_i = VAB - RA - OIN \quad (1)$$

siendo el VAB igual a la Producción (P) menos los Consumos intermedios (CI) de dicha rama o sector.

En el caso particular del sector hogares, el EBE se correspondería con el Excedente derivado de los servicios de alquiler de las viviendas ocupadas habitualmente por los hogares que son sus propietarios, que se conoce comúnmente como Alquiler imputado, incluido en la rama de actividad dedicada a las Actividades inmobiliarias. Por tanto, el EBE de los hogares

asociado a este Alquiler imputado tendría los componentes presentados en (1).

Cuadro 2.6
**CUENTAS DE RENTA DEL SECTOR HOGARES EN EL SEC 2010
COMPATIBLES CON LOS INGRESOS DE LA ECV**

II.1.2 CUENTA DE ASIGNACIÓN DE LA RENTA PRIMARIA PARA EL SECTOR HOGARES

<i>Empleos</i>	<i>Recursos</i>
D.4 Rentas de la propiedad	B.2b Excedente de explotación bruto/B.3b Renta mixta bruta
D.41 Intereses	B.2n Excedente de explotación neto/B.3n Renta mixta neta
D.45 Rentas de la tierra	D.1 Remuneración de asalariados
	D.11 Sueldos y salarios
	D.12 Cotizaciones sociales a cargo de los empleadores
	D.4 Rentas de la propiedad
	D.41 Intereses
	D.42 Rentas distribuidas de las sociedades
	D.44 Otras rentas de inversión
	D.45 Rentas de la tierra
B.5b Saldo de rentas primarias bruto	
B.5n Saldo de rentas primarias neto	

II.2 CUENTA DE DISTRIBUCIÓN SECUNDARIA DE LA RENTA PARA EL SECTOR HOGARES

<i>Empleos</i>	<i>Recursos</i>
D.5 Impuestos corrientes sobre la renta, el patrimonio, etc	B.5b Saldo de rentas primarias
D.61 Cotizaciones sociales netas	B.5n Saldo de rentas primarias neto
D.62 Prestaciones sociales distintas de las transferencias sociales en especie	D.5 Impuestos corrientes sobre la renta, el patrimonio, etc
D.7 Otras transferencias corrientes	D.61 Cotizaciones sociales netas
	D.62 Prestaciones sociales distintas de las transferencias sociales en especie
	D.7 Otras transferencias corrientes
B.6b Renta disponible bruta	
B.6n Renta disponible neta	

Fuente: SEC 2010

Según el SEC 2010, la producción por cuenta propia de servicios de alquiler de las viviendas ocupadas por los hogares que son sus propietarios, es decir, la producción de Alquiler imputado, se mide por el valor estimado del alquiler que pagaría un inquilino por el mismo alojamiento, tomando en consideración elementos como la situación, el equipamiento de la zona, etc., así como el tamaño y la calidad de la vivienda en cuestión³².

Los CI derivados de la producción de los servicios de las viviendas ocupadas por sus propietarios, que como para cualquier otra rama consisten en los bienes y servicios consumidos como insumos en un proceso productivo, se distribuyen, atendiendo a las Tablas de Origen y Destino de la economía española, entre los siguientes productos de la Clasificación Nacional de Productos por Actividades Año 2008 (CNPA-2008): *Trabajos de construcción*

³² El SEC 2010 contempla este mismo tipo de imputación para los garajes separados de las viviendas utilizados por sus propietarios como consumo final al mismo tiempo que éstas. Por su parte, no es preciso realizar ninguna estimación para los garajes utilizados por sus propietarios con el único fin de aparcarse cerca de su lugar de trabajo. El valor del alquiler de las viviendas ocupadas por sus propietarios en el extranjero, por ejemplo, viviendas de vacaciones, no debe registrarse como parte de la producción interior, sino como importación de servicios, y el excedente neto de exportación correspondiente como renta primaria recibida del resto del mundo. Un tratamiento análogo se aplica a las viviendas ocupadas por sus propietarios cuando éstos son no residentes. En el caso de las viviendas en régimen de multipropiedad, una parte del servicio de alojamiento deberá registrarse como tal.

especializados, Servicios de intermediación financiera y Servicios de seguros y planes de pensiones:

- Los CI en *Trabajos de construcción especializados* están relacionados con los gastos de las grandes mejoras o reformas llevadas a cabo por el hogar en la vivienda principal en propiedad.
- Los CI en *Servicios de intermediación financiera* se atribuyen principalmente a los costes que tienen que hacer frente los hogares por los créditos hipotecarios vinculados a la adquisición de dicha vivienda.
- Y los *Servicios de seguros y planes de pensiones* pagados por los hogares en relación a la vivienda principal se deben a los seguros contratados por éstos y relacionados con la compra de la misma (seguro de hogar, seguro de vida, etc.).

Por otro lado, por definición, en CNE no hay insumo de trabajo en la producción de Alquiler imputado, por lo que no se registra ningún empleo para este tipo de servicios.

Finalmente, los Otros impuestos netos sobre la producción asociados al Alquiler imputado comprenden todos aquellos impuestos netos de subvenciones que soportan los hogares por el hecho de poseer la vivienda principal en propiedad.

Una vez analizados los componentes del EBE de los hogares en CNE, pasamos a estudiar si podemos calcular esta partida contable con los microdatos de la ECV y cómo podemos aproximarla.

En relación al Alquiler imputado, la ECV nos proporciona la variable HY030N, denominada *Alquiler imputado*. Dicho concepto se refiere en la encuesta al valor que se imputará a todos los hogares que no declaren pagar un alquiler completo, ya sea porque son ocupantes propietarios, porque ocupan una vivienda alquilada a un precio inferior al de mercado o porque ocupan una vivienda a título gratuito. El alquiler imputado sólo se calculará en el caso de las viviendas (y demás locales asociados, como garajes) utilizadas por los hogares como residencia principal. El valor que se imputa es el equivalente al alquiler que se pagaría en el mercado por una vivienda similar a la ocupada, menos cualquier alquiler realmente abonado, es decir:

- Si la vivienda es en régimen de alquiler a precio inferior al de mercado, debe restarse el alquiler realmente pagado. Si además el hogar inquilino percibe una ayuda de la Administración o de una institución sin ánimo de lucro, debe restarse esta ayuda.
- Si el hogar disfruta de la vivienda en propiedad debe restarse cualquier ayuda percibida de la Administración o de una institución sin ánimo de lucro para hacer frente a los gastos de la misma, así como los gastos menores de mantenimiento o los gastos de renovación.

Por tanto, en base a esta definición de la ECV, podemos decir que la variable HY030N incluye la producción de Alquiler imputado en CNE relativa a los servicios de alquiler de las viviendas ocupadas por los hogares que son sus propietarios (que designaremos por medio de la variable AIVPP_P), menos los consumos intermedios asociados a la renovación de dichas viviendas (AIVPP_CI_TCE), denominados en CNE *Trabajos de construcción especializados*.

En concreto, respecto a la producción de Alquiler imputado, si comparamos en ambas fuentes sus definiciones, vemos que no coinciden exactamente. Mientras que en CNE el EBE recoge únicamente la producción de los servicios de alquiler de las viviendas ocupadas por sus propietarios, en la ECV se recoge tanto la producción de estos servicios relativos a viviendas de las que los hogares son ocupantes-propietarios, como los relativos a las viviendas alquiladas a un precio inferior al de mercado y a las viviendas ocupadas a título gratuito (por ejemplo, la vivienda facilitada por el empleador a sus asalariados, a título gratuito o a cambio de un alquiler reducido, como residencia principal del hogar, en concepto de remuneración en especie). En CNE los servicios de alquiler de viviendas suministradas a los hogares de forma gratuita se registran únicamente como Formación Bruta de Capital Fijo de las unidades institucionales propietarias de dichas viviendas. En el caso de que se presten servicios de alquiler de viviendas a los hogares a precios económicamente no significativos en CNE se registrarían como producción de *Servicios inmobiliarios* en la rama Actividades Inmobiliarias. En definitiva, para equilibrar estos dos conceptos de producción de alquiler imputado en la ECV y en CNE se ha de imponer sobre la variable Alquiler imputado de la ECV, HY030N, la condición de que el hogar posea la vivienda principal en propiedad (HH020=1 -Régimen de tenencia de la vivienda en propiedad- en ECV 2009, y HH021=1 o HH021=2 -Régimen de tenencia de la vivienda en propiedad sin o con hipoteca- en ECV 2012), y para recogerlo utilizaremos la variable AIVPP_P.

En relación al resto de CI de la rama Alquiler imputado de CNE, encontramos en la ECV los *Intereses pagados de préstamos hipotecarios (brutos)*, HY100G, asimilables a los CI en *Servicios de intermediación financiera* de CNE. Esta variable se refiere en la ECV al importe bruto total de los intereses hipotecarios, antes de deducir cualquier reducción fiscal por la vivienda principal del hogar durante el período de referencia de los ingresos. Quedan excluidos dentro de la variable HY100G:

- Cualquier otro pago simultáneo al pago de la cuota hipotecaria, como el seguro de hipoteca o el seguro de vivienda y contenido.
- Cualquier otro pago hipotecario de una vivienda que no sea la principal.
- Los pagos de otros préstamos adicionales para gastos de vivienda (reparaciones, restauración, mantenimiento, etc.) o de cualquier otro tipo no relacionados con la vivienda.
- El reembolso del principal o capital invertido.

Respecto a los Otros impuestos netos sobre la producción asociados al Alquiler imputado, no se encuentra información alguna en la encuesta.

Una vez analizado el significado de EBE de los hogares en CNE y en la ECV, podemos decir que el concepto de EBE que podemos construir con la encuesta representa tan sólo una parte de la partida de CNE, ya que únicamente se refiere a la producción de alquiler imputado y a los CI debidos a las grandes reformas llevadas a cabo en la vivienda y al pago de préstamos hipotecarios, quedando, por lo tanto, sin representación en la ECV los CI debidos a los *Servicios de seguros y planes de pensiones* y a los Otros impuestos netos sobre la producción. Para intentar subsanar esta ausencia de información en la encuesta trataremos de estimar estos componentes por medio de las variables AIVPP_CI_Seg y AIVPP_OIN, respectivamente.

Finalmente, como en la ECV no se estima el Consumo de Capital Fijo relacionado con la producción de los servicios de Alquiler imputado, para aproximarnos al concepto bruto del Excedente de Explotación de los hogares en CNE (partida B.2b), tendremos que estimar para la ECV el Consumo de Capital Fijo atribuible al Excedente de Explotación mediante la variable CCFEE.

RENTA MIXTA

Según el SEC 2010, en las empresas no constituidas en sociedad pertenecientes a los hogares el saldo contable de la cuenta de explotación se denomina Renta mixta y se refiere a los ingresos por el trabajo por cuenta propia. Esta renta debe su nombre a que contiene, por una parte, los beneficios obtenidos por las empresas no constituidas en sociedad pertenecientes a dicho sector y, por otra, la remuneración del trabajo por cuenta propia realizado por el propietario o los miembros de su familia.

En la ECV este saldo contable se corresponde con la *Renta de los trabajadores por cuenta propia*, constituida por los ingresos percibidos por las personas, para sí o en relación con los miembros de su familia, como resultado del ejercicio presente o pasado de una actividad por cuenta propia.

La ECV da la consideración de actividades por cuenta propia aquellas en las que la remuneración depende directamente de los beneficios (o posibles beneficios) derivados de los bienes y servicios producidos (el autoconsumo se considera parte de los beneficios). El trabajador por cuenta propia toma las decisiones operacionales que afectan a la empresa o bien delega la toma de tales decisiones, pero conserva la responsabilidad de la buena marcha de la empresa (en este contexto, la noción de «empresa» incluye las operaciones de una sola persona). La remuneración de las aficiones se considera actividad por cuenta propia.

La renta de los trabajadores por cuenta propia se desglosa en la ECV en:

- *Beneficios o pérdidas monetarios brutos de trabajadores por cuenta propia (incluidos derechos de propiedad intelectual)*, que se corresponde con la variable PY050G (o PY050N en términos netos).
- *Valor de los bienes producidos para autoconsumo*, variable PY070G/PY070N en términos brutos/netos en la ECV 2008 y HY170N, únicamente disponible en términos netos en la ECV 2012.

La variable *Beneficios o pérdidas monetarios brutos de trabajadores por cuenta propia* (incluidos derechos de propiedad intelectual) incluye:

- Las ganancias o pérdidas netas de explotación correspondientes a los propietarios de una empresa no constituida en sociedad que trabajan en ella, una vez deducidos los intereses de los préstamos de la empresa.
- Los derechos de propiedad intelectual percibidos en concepto de escritos, invenciones, etc., no incluidos en los beneficios y pérdidas de empresas no constituidas en sociedad.
- Ingresos por alquileres de edificios, vehículos, material y demás propiedades de empresa no incluidos en los beneficios y pérdidas de empresas no constituidas en sociedad, una vez deducidos los costes

derivados, como los intereses de préstamos asociados, las reparaciones y el mantenimiento y los gastos de seguros.

Sin embargo, estos beneficios excluyen:

- Los honorarios de directivos percibidos por los propietarios de empresas constituidas en sociedad (se incluyen en la variable PY010G, *Renta bruta monetaria o cuasi monetaria del asalariado*).
- Los dividendos procedentes de empresas constituidas en sociedad (se incluyen en la variable *Intereses, dividendos y ganancias de inversiones de capital en empresas no constituidas en sociedad* -HY090G-).
- Los beneficios procedentes de capital invertido en una empresa no constituida en sociedad en la que la persona no trabaja (socio capitalista) (incluidos en HY090G).
- El alquiler de terrenos y los ingresos procedentes de inquilinos o arrendatarios (se incluyen en la variable HY040G, *Renta procedente del alquiler de una propiedad o terreno*).
- Ingresos por alquileres de viviendas, habitaciones o terrenos no incluidos en los beneficios y pérdidas de las empresas no constituidas en sociedad (se incluyen también en HY040G).

El *Valor de los bienes producidos para autoconsumo* se refiere al valor de los alimentos y las bebidas producidos y consumidos dentro del mismo hogar, y se valora al precio de mercado de los bienes producidos, una vez deducidos todos los gastos soportados en la producción.

Sin embargo, el autoconsumo de la ECV excluye tanto la valoración de los servicios domésticos, como cualquier producto destinado a la venta y cualquier cantidad que el trabajador por cuenta propia retire de la empresa (estos valores se incluyen en *Beneficios o pérdidas monetarios brutos de trabajadores por cuenta propia [incluidos derechos de propiedad intelectual]*).

Teniendo en cuenta las especificaciones anteriores, para aproximarnos con la encuesta lo máximo posible al concepto de Renta Mixta que nos proporciona la CNE, consideraremos los siguientes componentes:

- La Renta Bruta Monetaria de los trabajadores por cuenta propia, que designaremos como RMCP_B, en la que tendremos en cuenta todos los ingresos monetarios de este colectivo. Este tipo de renta la estimaremos como suma de la variable PY050G, *Beneficios o pérdidas monetarios brutos de trabajadores por cuenta propia (incluidos derechos de propiedad intelectual)*, y de una parte de la *Renta bruta monetaria o cuasi-monetaria del asalariado*, variable PY010G. Ello se debe a que en el análisis de los microdatos de la ECV hay un subconjunto de registros cuya situación profesional es Empleadores o Empresarios sin asalariados o Trabajadores independientes que tienen uno o varios trabajos, no declaran trabajar por cuenta ajena y tienen rentas salariales monetarias, y que trataremos por tanto como renta monetaria de los trabajadores por cuenta propia.
- La Renta Bruta No Monetaria de los trabajadores por cuenta propia, RNMCP_B, en la que incluiremos todos los ingresos no monetarios de este colectivo. La estimaremos como suma de la variable PY070G en ECV 2009 o HY170N en ECV 2012, *Valor de los bienes producidos para autoconsumo*, y de una parte de la *Renta bruta no monetaria del asalariado*, variable

PY020G, para recoger, de forma análoga a como ocurre con la Renta Monetaria de los Trabajadores por cuenta propia aquellos registros correspondientes a Empleadores o Empresarios sin asalariados o Trabajadores independientes que tienen uno o varios trabajos, no declaran trabajar por cuenta ajena y poseen rentas salariales no monetarias.

- Para este colectivo también estimaremos las Cotizaciones sociales de los trabajadores por cuenta propia, CSEFCP, por la parte de rentas salariales que les atribuimos, tal y como venimos comentando, a través de la variable que recoge las *Cotizaciones sociales a cargo del empleador* (PY030G).
- Para calcular la Renta Mixta de CNE con la ECV también tendremos que incluir los Ingresos monetarios recibidos por el alquiler de propiedades (excluidos terrenos), IMALQPRO_B: En CNE, cuando un hogar alquila una propiedad o parte de ella se considera que realiza una actividad económica y, como tal, tendrá una producción, unos consumos intermedios, así como un beneficio por el desarrollo de la misma que irá incluido en la partida de Renta Mixta. Como ya dijimos anteriormente, este tipo de ingresos se encuentran recogidos en la ECV dentro de la variable *Renta bruta procedente del alquiler de una propiedad o terreno* (HY040G), y, por tanto, tendremos que estimar en la encuesta las rentas que obtienen los hogares exclusivamente por alquileres de terrenos de las cuales son sus propietarios -para ello usaremos la variable IMALQTER_B-.
- Y, por último, dado que la ECV no incluye una estimación para el Consumo de Capital Fijo (CCF) de la Renta Mixta, para lograr una estimación lo más homogénea posible de esta partida en términos brutos de CNE (B.3b) en la ECV, tendremos que estimar también en ésta el CCF correspondiente a esta partida. Para ello, utilizaremos la variable CCFRM, que denominaremos Consumo de Capital Fijo atribuible a la Renta Mixta.

REMUNERACIÓN DE ASALARIADOS

La Remuneración de asalariados (D.1) comprende, en el SEC 2010, toda la remuneración en efectivo y en especie a pagar por los empleadores a sus asalariados como contrapartida del trabajo realizado por éstos durante el período contable. D.1 se desglosa en CNE en:

- a) Sueldos y salarios (D.11):
 - 1. Sueldos y salarios en efectivo.
 - 2. Sueldos y salarios en especie.
- b) Cotizaciones sociales a cargo de los empleadores (D.12):
 - 1. Cotizaciones sociales efectivas a cargo de los empleadores (D.121).
 - 2. Cotizaciones sociales imputadas a cargo de los empleadores (D.122).

Por su parte, la partida contable D.1 se corresponde en la ECV con la *Renta del Asalariado*, entendiéndose por tal la remuneración total, monetaria o no monetaria, que debe pagar un empleador a su asalariado como contrapartida del trabajo realizado por éste durante el período de referencia de los ingresos.

La *Renta del asalariado* de la ECV se desglosa en:

- a) Renta bruta monetaria o cuasi-monetaria del asalariado.
- b) Renta bruta no monetaria del asalariado.
- c) Cotizaciones sociales a cargo del empleador.

A continuación, analizaremos las definiciones de cada uno de estos conceptos en ambas fuentes con el objeto de estimar la Remuneración de asalariados en la ECV.

Sueldos y salarios en efectivo

Los sueldos y salarios en efectivo incluyen, según el SEC 2010, el valor de todas las cotizaciones sociales, impuestos sobre la renta, etc., a cargo del asalariado, incluso cuando el empleador los retiene y paga directamente a los sistemas de seguros sociales, las autoridades fiscales, etc., en nombre del asalariado.

En concreto, los Sueldos y salarios en efectivo incluyen los tipos de remuneración siguientes:

- Sueldos y salarios básicos pagados a intervalos regulares.
- Gratificaciones por horas extraordinarias, trabajo nocturno o en fin de semana, o trabajo en circunstancias desagradables o peligrosas.
- Complementos por carestía de vida, residencia y traslado al extranjero.
- Primas de productividad o sobre los beneficios, gratificaciones de Navidad y de fin de año, excluidas las prestaciones sociales directas de los empleadores y las pagas extraordinarias anuales.
- Complementos de transporte al y desde el lugar de trabajo, excluidos los complementos o reembolsos a los asalariados por gastos de viaje, alojamiento, mudanza y representación en el ejercicio de sus funciones.
- Pagos por días festivos no laborables y vacaciones anuales.
- Comisiones, propinas, primas de asistencia y gratificaciones pagadas a los asalariados.
- Primas especiales u otros pagos excepcionales vinculados al rendimiento global de la empresa dentro de programas de incentivos.
- Pagos efectuados por los empleadores a sus asalariados en el marco de planes de ahorro.
- Pagos excepcionales a los asalariados que abandonen la empresa, si no están vinculados a un convenio colectivo.
- Ayudas en efectivo para alquiler de viviendas pagadas por los empleadores a sus asalariados.

Por su parte, en la ECV los Sueldos y salarios en efectivo se recogen con el nombre de *Renta Bruta Monetaria o Cuasi monetaria del asalariado* (PY010G) y se refieren a la componente monetaria de la remuneración que el empleador debe pagar a sus asalariados. Incluye el valor de todas las cotizaciones sociales e impuestos sobre la renta que el asalariado (o el empleador en nombre del asalariado) debe pagar a los sistemas de seguros sociales o a las autoridades fiscales.

La *Renta bruta monetaria o cuasi-monetaria del asalariado* incluye los siguientes elementos:

- Sueldos y salarios monetarios por el tiempo trabajado o por el trabajo realizado en la actividad principal o en cualquier otra actividad secundaria u ocasional.
- Remuneración del tiempo no trabajado (por ejemplo, vacaciones anuales).
- Gratificación por horas extraordinarias.
- Honorarios destinados a los directivos de empresas constituidas en sociedad.
- Remuneración por trabajo a destajo.
- Comisiones, propinas y gratificaciones.
- Pagos adicionales (por ejemplo, paga extraordinaria).
- Participación en los beneficios y primas monetarias.
- Primas de productividad.
- Complementos por trabajar en lugares remotos (se considera parte de las condiciones del puesto de trabajo).
- Complementos de transporte hasta el lugar de trabajo y desde éste.

Mientras que PY010G excluye:

- Los reembolsos por parte del empleador de los gastos relacionados con el trabajo (desplazamientos por trabajo, por ejemplo).
- Las indemnizaciones por cese destinadas a compensar a los asalariados por la interrupción de su trabajo antes de haber alcanzado la edad normal de jubilación para el trabajo desempeñado, así como las indemnizaciones por despido (se incluyen en *Prestaciones por desempleo* -PY090G-).
- Las ayudas para gastos puramente laborales, como los gastos de viaje y dietas o la compra de ropa de protección.
- Los pagos a tanto alzado en la fecha normal de jubilación (se incluyen en *Prestaciones por jubilación* -PY100G-).
- Los subsidios de huelga.

Así, a la vista de las similitudes en las definiciones de CNE y ECV, la partida contable Sueldos y Salarios en efectivo, que vamos a calcular en la ECV por medio de la variable Renta Monetaria por Cuenta Ajena (RMCA_B), la haremos corresponder con la variable PY010G, agregada para todas las personas encuestadas, a excepción de la parte de esta variable considerada como *Renta Monetaria de los trabajadores por cuenta propia* que, como ya vimos está incluida dentro de la Renta Mixta.

Adicionalmente, en CNE, dentro de la partida Sueldos y Salarios en efectivo se incluyen las cotizaciones sociales efectivas pagadas por los asalariados (D.6112). En la ECV, estas cotizaciones se incluyen de forma agregada en la variable HY140G, junto a otro tipo de cotizaciones y al Impuesto sobre la renta. Pero, a su vez, estas cotizaciones las incluyen todas las variables de la encuesta expresadas en términos brutos que son con las que estamos trabajando, razón por la cual no tenemos que volver a sumarla.

Sueldos y salarios en especie

Por su parte, los Sueldos y salarios en especie en el SEC 2010 comprenden los bienes y servicios, u otras prestaciones, proporcionados gratuitamente o a precios reducidos por los empleadores y que los asalariados pueden utilizar en su tiempo libre y de la forma que deseen para la satisfacción de sus necesidades o las de otros miembros de sus hogares. Estos bienes y servicios,

así como las otras prestaciones, no son necesarios para el proceso de producción del empleador. Para los asalariados, estos sueldos y salarios en especie representan una renta adicional, ya que, de haberlos comprado ellos mismos, los habrían pagado al precio de mercado. Los más comunes son:

- Comidas y bebidas consumidas en los viajes de negocios y las reducciones de precios obtenidos en los comedores de empresa gratuitos o subvencionados, o con los bonos de comida.
- Servicios de alquiler de viviendas (servicios de alojamiento) que han sido adquiridos o producidos por cuenta propia y que pueden ser utilizados por todos los miembros del hogar del asalariado.
- Uniformes y otros tipos de ropa de trabajo especial que los asalariados optan por llevar tanto en el trabajo como fuera de éste.
- Servicios de vehículos automóviles y de otros bienes duraderos facilitados para el uso personal de los asalariados.
- Bienes y servicios obtenidos en los propios procesos de producción, tales como viajes gratuitos de los asalariados de las compañías de ferrocarriles o aéreas, carbón gratuito para los mineros o alimentos gratuitos para los asalariados de la agricultura.
- Servicios de las instalaciones deportivas, recreativas o de vacaciones que están a disposición de los asalariados y sus familias.
- Transporte al y desde el lugar de trabajo, excepto cuando se efectúa durante la jornada de trabajo, y el servicio de aparcamiento.
- Servicios de guardería para los hijos de los asalariados.
- Pagos efectuados por los empleadores a comités de empresa u órganos similares.
- Acciones gratuitas distribuidas a los asalariados.
- Préstamos a los asalariados a tipos de interés reducidos; el valor de esta prestación se calcula como la diferencia entre el importe que debería pagar el asalariado si se le cargara el tipo de interés en condiciones de mercado y el importe realmente pagado; la prestación se registra en los sueldos y salarios de la cuenta de explotación, y el correspondiente pago de intereses imputados del asalariado al empleador se registra en las cuentas de distribución primaria de la renta.
- Opciones de compra de acciones, cuando un empleador ofrece al asalariado la opción de comprar acciones o participaciones a un precio determinado en una fecha futura.
- Rentas generadas por actividades no observadas³³ en sectores empresariales y transferidas a los asalariados que participan en dichas actividades para su uso privado.

En la ECV, los Sueldos y Salarios en especie se recogen en la variable *Renta Bruta No monetaria del asalariado*, PY020G, que se refiere a las componentes no monetarias de la renta que el empleador puede proporcionar al

³³ Las actividades no observadas se refieren a aquellas actividades de producción (que a su vez generan actividades de intercambio y consumo) que, por razones fiscales o de índole legal permanecen sin control de la administración pública. Obviamente, todas estas actividades generan producción y riqueza. Este tipo de actividades se engloba dentro de la Economía No Observada (ENO) que comprende, a su vez, la economía sumergida, la economía ilegal, la economía informal y la producción propia de los hogares. La entrada en vigor del SEC 2010 ha hecho que los países de la Unión tengan que hacer un esfuerzo especial para incorporar estas actividades no observadas en sus cuentas con el fin de reducir la minusvaloración del PIB por estos conceptos. En particular, se deben incluir obligatoriamente las actividades ilegales de contrabando de tabaco, juego ilegal, prostitución organizada y tráfico de drogas.

asalariado, a título gratuito o a precio reducido, como parte de las ventajas del contrato. La variable PY020G incluye:

- El vehículo de empresa, tanto para uso privado sólo como para uso privado y profesional, y los costes derivados del uso privado (combustible gratuito, seguro, tasas e impuestos aplicables, etc.).
- Las comidas gratuitas o subvencionadas y los cheques restaurante.
- El reembolso o abono de los gastos relacionados con la vivienda (gas, electricidad, agua, teléfono o facturas del teléfono móvil, entre otros).
- Otros bienes y servicios que proporciona el empleador a sus asalariados, a título gratuito o a precio reducido, cuando constituyen una componente significativa de la renta en el ámbito nacional o una componente significativa de la renta de grupos determinados de hogares.

El valor de los bienes y servicios proporcionados a título gratuito en la ECV se calculará con arreglo al precio de mercado de dichos bienes y servicios. El valor de los bienes y servicios proporcionados a precio reducido corresponderá a la diferencia entre el precio de mercado y la cantidad abonada por el asalariado.

Por su parte, en la ECV se excluyen:

- El coste soportado por el empleador en concepto de suministro de cualquiera de dichos bienes y servicios, si éstos sólo son necesarios para el desempeño del trabajo por parte de los asalariados.
- Los servicios de vivienda en el lugar de trabajo si éstos no pueden ser utilizados por los hogares a los que pertenecen los asalariados.
- La vivienda facilitada por el empleador a sus asalariados, a título gratuito o a cambio de un alquiler reducido, como residencia principal del hogar (el valor imputado de la vivienda facilitado a título gratuito o a cambio de un alquiler reducido se incluye en *Alquiler imputado*).
- La vivienda facilitada por el empleador a sus asalariados, a título gratuito o a cambio de un alquiler reducido, como residencia secundaria del hogar.
- Las ayudas concedidas a los asalariados para la compra de herramientas, material, ropa, etc., necesarios exclusiva o principalmente para el desempeño de su trabajo.
- Los alimentos o bebidas especiales necesarios en condiciones de trabajo excepcionales.
- Cualquier bien o servicio que perciban los asalariados en el lugar de trabajo o que necesiten debido a la naturaleza de la actividad (revisión médica necesaria para trabajar, por ejemplo).

Una vez estudiadas las definiciones en ambas fuentes, cabe destacar que, si bien presentan algunas discrepancias, en general, son bastante próximas entre sí. Una de las diferencias más importantes es el tratamiento de los servicios de alquiler de vivienda facilitada por el empleador a sus asalariados, como residencia principal del hogar. Así, mientras que en el SEC 2010 se tratan como Sueldos y Salarios en especie, en la ECV se incluyen en la variable *Alquiler imputado*.

Al margen de estos matices conceptuales, el total de Sueldos y Salarios en especie de Contabilidad Nacional, lo vamos a representar en la ECV por medio de la variable RNMCA_B, *Renta No Monetaria por Cuenta Ajena*, y para su cálculo utilizaremos la variable agregada PY020G, *Renta bruta no*

monetaria del asalariado, salvo la parte de ésta atribuible a la Renta No Monetaria de los trabajadores por cuenta propia, que ya mencionamos en la explicación de la Renta Mixta.

Para terminar de estimar con la encuesta la variable Sueldos y Salarios de CNE, tendremos en cuenta también una parte de la variable de la ECV denominada *Renta bruta percibida por los menores de 16 años* (HY110G), que incluye, entre otras, las Rentas brutas del trabajo por cuenta ajena percibida por los menores de 16 años, tanto monetarias como no monetarias. Esto es debido a que, según el Estatuto de los Trabajadores publicado en el Boletín Oficial del Estado el 10 de marzo de 1980 y modificado por el Real Decreto Legislativo 1/1995, de 24 de marzo, por el que se aprueba el texto refundido de la Ley del Estatuto de los Trabajadores, los menores de 16 años pueden ser contratados, previa autorización de los padres o tutores y del organismo competente. Así, se les permite participar en espectáculos públicos siempre y cuando su intervención no suponga un peligro para su integridad física y psicológica. Para la estimación del conjunto de Rentas del trabajo por cuenta ajena recibidas por los menores de 16 años creamos la variable ORNC1_B.

Por tanto, la rúbrica contable D.11, Sueldos y Salarios, se obtendrá en la ECV como suma de las variables PY010G y PY020G agregadas para todas las personas encuestadas, a excepción de la parte de éstas correspondiente a los trabajadores por cuenta propia, más la variable ORNC1_B referida a las rentas del trabajo percibidas por los menores de 16 años.

Para la mayor comprensión de las siguientes partidas, conviene aclarar que en CNE no se consideran Sueldos y salarios:

- a) Los gastos de los empleadores necesarios para su proceso de producción (por ejemplo, los complementos o reembolsos a los asalariados por gastos de viaje, traslado, mudanza y representación en el ejercicio de sus funciones; los gastos relativos al acondicionamiento del lugar de trabajo, los exámenes médicos necesarios por el tipo de trabajo y el suministro de ropa de trabajo que se utilice con tal fin; etc.). Estos gastos en bienes y servicios se consideran consumos intermedios de los empleadores.
- b) Los importes de los sueldos y salarios que los empleadores pagan temporalmente a sus asalariados en caso de enfermedad, maternidad, accidente de trabajo, invalidez, etc. Dichos pagos se consideran otras *Prestaciones sociales distintas de las pensiones* (D.6222) y figuran por el mismo importe en las *Cotizaciones sociales imputadas a cargo de los empleadores distintas de las pensiones* (D.1222).
- c) Las otras prestaciones de seguros sociales vinculadas con el empleo, en forma de complementos por hijos o cónyuges y complementos familiares, de educación o de otro tipo en relación con las personas a cargo del asalariado y en forma de servicios médicos gratuitos (excluidos los necesarios por la naturaleza del trabajo) para los asalariados y sus familias.
- d) Cualquier impuesto que deba pagar el empleador sobre los sueldos y salarios (por ejemplo, un impuesto sobre la masa salarial). Estos impuestos se consideran impuestos sobre la producción.
- e) Los pagos a los trabajadores a domicilio según tarifas por pieza. Cuando la renta que recibe un trabajador a domicilio depende del resultado de algún proceso de producción del que esa persona es responsable,

independientemente de cuánto trabajo se haya o no realizado, este tipo de remuneración supone que el trabajador es autónomo.

Cotizaciones sociales a cargo de los empleadores

A continuación, pasamos a estudiar la otra gran partida contable de la *Remuneración de asalariados*, las *Cotizaciones sociales a cargo de los empleadores*.

Según el SEC 2010, en la *Remuneración de los asalariados* debe registrarse un montante igual al importe de las *Cotizaciones sociales a cargo de los empleadores* (D.12) para garantizar a sus asalariados el derecho a percibir prestaciones sociales. Dichas cotizaciones pueden ser efectivas o imputadas.

Las *Cotizaciones sociales efectivas a cargo de los empleadores* (D.121) comprenden los pagos que éstos realizan en beneficio de sus asalariados a las entidades aseguradoras (seguridad social y otros sistemas de seguros sociales vinculados al empleo). Estos pagos cubren a la vez las cotizaciones legales, contractuales y voluntarias de los seguros contra riesgos y necesidades sociales. Aunque los empleadores las pagan directamente a las entidades aseguradoras, se considera que estas cotizaciones son un componente de la remuneración de los asalariados, y se registra que son éstos quienes las pagan a las entidades aseguradoras. En concreto este tipo de cotizaciones comprenden las cotizaciones vinculadas a las pensiones y las cotizaciones para otras prestaciones sociales distintas de las pensiones, como las de enfermedad, maternidad, accidente laboral, incapacidad, despido, etc.

Por su parte, las *Cotizaciones sociales imputadas a cargo de los empleadores* (D.122) representan la contrapartida de Prestaciones de otros sistemas de seguros sociales (D.622) (menos, en su caso, las cotizaciones sociales a cargo de los asalariados). Dichas prestaciones las pagan directamente los empleadores a sus asalariados o antiguos asalariados y otros derechohabientes sin la intervención de una empresa de seguro o un fondo de pensiones autónomo, y sin constituir un fondo especial o una reserva separada a tales efectos. Se pagan por pensiones y en casos de enfermedad, maternidad, accidente de trabajo, incapacidad, despido, etc. El hecho de que estas prestaciones sociales sean pagadas directamente por los empleadores y no por medio de administraciones de seguridad social o de otras entidades aseguradoras no desvirtúa su carácter de prestaciones sociales. No obstante, como los costes de estas prestaciones constituyen parte de los costes salariales del empleador, se deberán incluir también en la remuneración de los asalariados.

El importe de las *Cotizaciones sociales imputadas a cargo de los empleadores* tiene en cuenta las cotizaciones efectivas a cargo del empleador o del asalariado y depende no solo de las cuantías de las prestaciones a pagar en el período corriente, sino también de la forma en que probablemente evolucionen en el futuro las obligaciones de los empleadores bajo esos sistemas como resultado de factores tales como las variaciones previstas del número de empleados, la pirámide de edad y la esperanza de vida de sus asalariados y antiguos asalariados. Los valores imputados de estas cotizaciones se basan en el mismo tipo de cálculos actuariales que determinan los niveles de las primas cobradas por las compañías de seguros.

En lo que se refiere a la ECV, dentro de la Renta del Asalariado, la variable *Cotizaciones sociales brutas a cargo del empleador*, (PY030G), comprende los pagos realizados por éste durante el periodo de referencia de los ingresos a las entidades aseguradoras (sistema de la de seguridad social y sistemas de fondos privados) en beneficio de sus asalariados; se consideran las cotizaciones legales, contractuales y convencionales en concepto de protección contra los riesgos sociales. Se incluyen los siguientes conceptos:

- Cotizaciones del empleador a sistemas de protección social públicos (seguridad social).
- Cotizaciones del empleador a planes de jubilación privados (pensiones).
- Cotizaciones del empleador a planes de seguro de enfermedad privados.
- Cotizaciones del empleador a planes de seguro de vida.
- Cotizaciones del empleador a otros sistemas de seguros (por ejemplo, invalidez).

Comparando las definiciones de cotizaciones en ambas fuentes podemos ver que, en la ECV se recogen las que pagan los empleadores a la seguridad social u a otras formas aseguramiento, y no las que paga la propia empresa sin que haya ninguna forma de aseguramiento de por medio, por lo que se corresponden con las cotizaciones sociales efectivas del SEC 2010. No encontramos en la ECV ninguna variable que recoja información sobre las *Cotizaciones sociales imputadas a cargo de los empleadores*.

En definitiva, para calcular en la ECV las *Cotizaciones sociales a cargo de los empleadores* registradas en el SEC 2010 bajo la rúbrica D.12, utilizaremos, por una parte, la variable PY030G, *Cotizaciones sociales a cargo del empleador*, para estimar las Cotizaciones sociales efectivas a cargo del empleador, detrayendo de éstas las *Cotizaciones sociales efectivas de los trabajadores por cuenta propia* que ya explicamos que se incluían en la *Renta Mixta*, CSEFCP, y adicionándole una estimación de las *Cotizaciones sociales imputadas a cargo de los empleadores*, para lo que utilizaremos la variable CSIMEMP.

RENTAS DE LA PROPIEDAD

En este apartado, estudiaremos la homogeneidad del conjunto de *Rentas de la propiedad* referidas al sector hogares de la ECV respecto a la Contabilidad Nacional.

A priori podemos decir que la definición del concepto general de Rentas de la propiedad en una y otra fuente de información es idéntica, esto es, rentas que recibe el propietario –en nuestro estudio, el hogar- de un activo financiero o de un activo material no producido a cambio de proporcionar fondos o de poner el activo material no producido a disposición de otra unidad institucional. No obstante, a continuación, veremos que, a pesar de la similitud del concepto general en ambas fuentes, encontramos alguna diferencia en su composición.

El SEC 2010 distingue, para el sector hogares, entre las siguientes *Rentas de propiedad* (D.4): *Intereses*, *Rentas distribuidas de las sociedades*, *Otras rentas de inversión* y *Rentas de la tierra*. Veamos cómo se definen cada una de ellas en nuestro marco contable de referencia y cómo se recogen en la ECV.

Intereses

El SEC 2010 define la partida Intereses (D.41) como las rentas de propiedad que reciben los propietarios de activos financieros por ponerlos a disposición de otra unidad institucional por medio de los siguientes instrumentos financieros: *efectivo y depósitos, valores representativos de deuda, préstamos y otras cuentas pendientes de cobro/pago*.

Los activos financieros que generan intereses son los derechos de los acreedores sobre los deudores. Los acreedores prestan fondos a los deudores que dan lugar a la creación de alguno de los instrumentos financieros antes enumerados.

Se diferencian varios tipos de intereses:

- 1) **Intereses de depósitos y préstamos.** El importe de los intereses de depósitos y préstamos a cobrar y a pagar de las instituciones financieras incluye un ajuste del margen que representa un pago implícito por los servicios prestados por las instituciones financieras en la concesión de préstamos y la aceptación de depósitos. El pago o ingreso se divide entre la parte del servicio y la que corresponde al concepto de interés en las cuentas nacionales. Los pagos o ingresos efectivos recibidos de instituciones financieras o abonados a estas, denominados **intereses bancarios**, deben dividirse de modo que el concepto de interés de las cuentas nacionales y los gastos del servicio se registren por separado. La diferencia negativa entre el importe de los intereses en las cuentas nacionales pagados por los prestatarios a las instituciones financieras y los intereses bancarios corresponde al valor estimado de los gastos a pagar, mientras que la diferencia positiva entre el importe de los intereses en las cuentas nacionales a cobrar por los depositantes y el interés bancario representa el importe del servicio a pagar. Los valores de los gastos se registran como ventas de servicios en la cuenta de producción de las instituciones financieras y como empleos en la cuenta de sus clientes.
- 2) **Intereses de los valores representativos de deuda.** Incluyen los intereses de las letras e instrumentos similares a corto plazo, así como los intereses de los bonos y obligaciones.
- 3) **Permutas financieras (swaps) de tipos de interés y acuerdos de tipos de interés futuros (forward rate agreements).** Los pagos resultantes de cualquier tipo de acuerdo de permuta financiera (swap) no se considerarán intereses ni se registrarán como rentas de la propiedad, sino como operaciones en los derivados financieros de la cuenta financiera. Igualmente, las operaciones mediante acuerdos de tipos de interés futuros se registran como operaciones en los derivados financieros de la cuenta financiera, y no como rentas de la propiedad.
- 4) **Intereses de arrendamientos financieros.** Un arrendamiento financiero es una forma de financiar, por ejemplo, la compra de maquinaria y bienes de equipo. El arrendador compra el equipo y el arrendatario se compromete contractualmente a pagar, durante el período del contrato, unos alquileres con los que el arrendador recupera los costes, incluidos los intereses a los que renuncia por el dinero utilizado en adquirir el equipo.

Se considera que el arrendador hace un préstamo al arrendatario igual al valor del precio de adquisición del activo, préstamo que se va devolviendo gradualmente durante el período de arrendamiento. Por lo tanto, se considera que la renta que paga el arrendatario en cada período tiene dos componentes: la amortización del principal y el pago de los intereses. El tipo de interés del préstamo imputado se determina mediante la relación entre el importe total pagado en concepto de alquiler durante el arrendamiento y el precio de adquisición del activo. La parte del alquiler que representa los intereses disminuye gradualmente durante la duración del arrendamiento, conforme se va amortizando el principal. El préstamo inicial al arrendatario, junto con el subsiguiente reembolso del principal, se registran en las cuentas financieras del arrendador y del arrendatario. Los pagos de intereses se registran como intereses en la cuenta de distribución primaria de la renta.

- 5) **Otros intereses.** Se tratan también como intereses: las comisiones bancarias por descubierto; las primas de fidelidad; y los pagos determinados mediante sorteo a los tenedores de bonos.

Por el lado de la ECV, y con menor nivel de detalle, encontramos que la variable mediante la que podemos aproximar los *Intereses* de CNE es la denominada *Intereses, dividendos y ganancias brutos de inversiones de capital en empresas no constituidas en sociedad* (HY090G). En concreto, esta componente de la encuesta recoge los intereses procedentes de los activos como cuentas bancarias, certificados de depósito, bonos, etc. y, además, los dividendos y las ganancias procedentes de inversiones de capital en una empresa no constituida en sociedad, en la cual no trabaja la persona (socio capitalista), percibidos durante el período de referencia de los ingresos, una vez deducidos los gastos soportados.

En base a esta definición podemos decir que, a diferencia de CNE que distingue entre Intereses recibidos y pagados, la ECV recoge únicamente los *Intereses* recibidos por los hogares -equiparables a los recursos de la partida contable D.41, D.41_r. Además, éstos se recogen en la encuesta junto a los dividendos y ganancias brutos de inversiones de capital en empresas no constituidas en sociedad, los cuales se estiman en CNE bajo la rúbrica D.42. que pasamos a estudiar a continuación.

Rentas distribuidas de las sociedades

El SEC 2010 distingue dos tipos de rentas dentro de esta partida: los *Dividendos* y las *Rentas retiradas de las cuasi-sociedades*.

Los *Dividendos* (D.421) son una forma de rentas de la propiedad a la que tienen derecho los propietarios de *Participaciones en el capital y en fondos de inversión*, como resultado, por ejemplo, de haber puesto fondos a disposición de las sociedades.

La captación de capital mediante la emisión de acciones constituye una alternativa a la captación de fondos mediante préstamos. No obstante, al contrario de lo que ocurre con el capital prestado, el capital de las acciones no da lugar a un pasivo fijo en términos monetarios ni da derecho a los tenedores de las acciones de una sociedad a recibir una renta fija o

predeterminada. Se entiende por dividendos toda distribución de beneficios por las sociedades a sus accionistas o propietarios.

En CNE, dentro de la rúbrica *Dividendos* se incluyen también:

a) las acciones distribuidas a los accionistas como pago del dividendo del ejercicio financiero;

b) la renta pagada a las administraciones públicas por las empresas públicas dotadas de personalidad jurídica, que no están constituidas formalmente en sociedad;

c) las rentas generadas por actividades no observadas y transferidas a los propietarios de las sociedades que participan en dichas actividades para su uso privado.

Por su parte, las *Rentas retiradas de las cuasi-sociedades* (D.422) comprenden los importes que los empresarios retiran efectivamente, para su propio uso, de los beneficios realizados por las cuasi-sociedades que les pertenecen. Estos importes se registrarán antes de la deducción de los impuestos corrientes sobre la renta, el patrimonio, etc., que se consideran siempre a pagar por los propietarios de las empresas.

Según CNE, cuando una cuasi-sociedad obtiene beneficios comerciales, la unidad propietaria puede dejarlos a disposición de la cuasi-sociedad, en parte o en su totalidad, fundamentalmente con fines de inversión. Esta renta que se deja a disposición de las cuasi-sociedades aparece como un ahorro de estas últimas y solamente los beneficios realmente retirados por las unidades propietarias se contabilizan como *Rentas retiradas de las cuasi-sociedades*.

La rúbrica *Rentas retiradas de las cuasi-sociedades* no incluye los importes que los propietarios reciben en concepto de: venta de bienes de capital fijo existentes; venta de terrenos y activos no producidos; y retiradas de capital (por ejemplo, la liquidación total o parcial de su participación en la cuasi-sociedad). Estos importes se registran en la cuenta financiera como disminuciones de la participación de los propietarios.

En definitiva, teniendo en cuenta la disponibilidad de información en la ECV, los *Intereses* recibidos publicados en la CNE, D41_r, junto con los de esta última rúbrica que acabamos de definir, D.42_r, los podemos aproximar en la ECV a la variable HY090G, *Intereses, dividendos y ganancias brutos de inversiones de capital en empresas no constituidas en sociedad*. A lo largo de este trabajo utilizaremos la variable IMCAP_B para referirnos a HY090G. Por su parte, los *Intereses* pagados de CNE, D.41_e, no recogidos en la ECV, tendremos que estimarlos y, para ello, utilizaremos la variable INTP_B.

Otras rentas de inversión

En el Sistema SEC 2010, las *Otras rentas de inversión* (D.44) se dividen en: *Rentas de inversión atribuibles a los tomadores del seguro* (D.441), *Rentas de inversión a pagar sobre los derechos por pensiones* (D.442) y *Rentas de inversión atribuibles a partícipes en fondos de inversión* (D.443).

Las *Rentas de inversión atribuibles a los tomadores del seguro* (D.441) corresponden al total de las rentas primarias recibidas de la inversión de las

reservas técnicas de seguro. Las reservas son aquellas en las que la empresa de seguro reconoce un pasivo correspondiente con respecto a los tomadores del seguro. Las compañías de seguros invierten sus reservas técnicas en activos financieros o en terrenos (de los que provienen las rentas de la propiedad netas, es decir, una vez deducidos los intereses pagados) o en inmuebles (que generan excedentes de explotación netos).

Las rentas D.441 se registran por separado entre los tomadores de pólizas de seguro de vida y los tomadores de pólizas de seguro no vida.

- En el caso de las pólizas no vida, la empresa de seguro tiene un pasivo con respecto al tomador del seguro equivalente al importe de la prima depositada en la empresa, pero aún no consumida, al valor de las indemnizaciones imputables al ejercicio aún no abonadas y a una reserva para las indemnizaciones todavía no notificadas o notificadas, pero todavía sin liquidar. La empresa de seguro dispone de reservas técnicas para hacer frente a este pasivo. Las rentas de inversión de estas reservas técnicas se consideran rentas atribuibles a los tomadores del seguro, distribuidas a continuación a estos en la cuenta de asignación de la renta primaria y reembolsadas a la empresa de seguro como prima complementaria en la cuenta de distribución secundaria de la renta.
- En el caso de las pólizas y rentas de seguros de vida, las compañías de seguros disponen de pasivos con respecto a los tomadores de las pólizas y los rentistas, equivalentes al valor actual de las indemnizaciones previstas. Para hacer frente a estos pasivos, las compañías de seguros disponen de fondos que pertenecen a los tomadores de las pólizas y que consisten en bonos declarados para las pólizas con beneficios, así como disposiciones para los tomadores de las pólizas y los rentistas para el pago de futuros bonos y otras indemnizaciones. Dichos fondos se invierten en diferentes activos financieros y no financieros.

Las *Rentas de inversión a pagar sobre los derechos por pensiones* (D.442) se derivan de los dos tipos de sistemas de pensiones, a saber, los sistemas de cotización definida y los sistemas de prestación definida.

Un sistema de cotización definida es aquel en el que las cotizaciones tanto de los empleadores como de los asalariados se invierten en nombre de estos últimos como futuros pensionistas. No existe ninguna otra fuente de financiación de las pensiones y los fondos no se utilizan con ningún otro fin. Las rentas de inversión a pagar por los derechos de cotización definida equivalen a las rentas de inversión de los fondos más todo ingreso obtenido del alquiler de tierras o bienes inmuebles propiedad del fondo.

La característica de un sistema de prestación definida es que el nivel de los pagos a los pensionistas se determina a partir de una fórmula, lo que hace posible establecer el nivel de los derechos como valor actual de los futuros pagos, calculado utilizando hipótesis actuariales sobre la duración de vida e hipótesis económicas sobre los tipos de interés o la tasa de descuento. El valor actual de los derechos existentes al comienzo del año aumenta, ya que la fecha en la que deben abonarse dichos derechos está un año más próxima. Este aumento se considera renta de inversión atribuida a los titulares de pensiones en el caso del sistema de prestación definida. El importe del

aumento no se ve afectado por el hecho de que el sistema de pensiones tenga o no realmente fondos suficientes para atender a todas sus obligaciones ni por el tipo de incremento de los fondos, ya sean rentas de inversión o ganancias de posesión.

Las *Rentas de inversión atribuibles a partícipes en fondos de inversión* (D.443), incluidos los fondos de inversión y las sociedades de inversión se componen de las rúbricas *Dividendos atribuibles a partícipes en fondos de inversión* (D.4431) y *Beneficios no distribuidos atribuibles a partícipes en fondos de inversión* (D.4432). Este tipo de rentas de la propiedad recibidas por los fondos de inversión se registran como rentas de la propiedad de los partícipes en el fondo aun cuando no se distribuyen, sino que se reinvierten en su nombre.

Respecto de la otra fuente de información que venimos analizando de forma paralela al SEC 2010, tenemos que decir que las *Otras Rentas de Inversión* no se recogen en la ECV, razón por la cual, procederemos a su estimación por medio de la variable ORI_B.

Rentas de la tierra

En CNE dentro de esta partida (D.45) se incluyen a su vez dos tipos de rentas: las *Rentas de terrenos* y las *Rentas de los activos del subsuelo*, aunque desconocemos este desglose cuantitativamente hablando ya que en la publicación de CNE no se desglosa el montante correspondiente a cada uno de ellas.

En cuanto a las *Rentas de terrenos* hay que decir que son aquellas que recibe un propietario de terrenos de un arrendatario. Esta rúbrica comprende también las rentas a pagar a los propietarios de ríos, lagos, lagunas y otras superficies de agua por el derecho a explotarlos con fines recreativos o de otro tipo, incluida la pesca.

Un propietario puede estar obligado a pagar impuestos sobre los terrenos o incurrir en determinados gastos de mantenimiento por el mero hecho de ser propietario de los mismos. Por convenio, estos impuestos y gastos se consideran a pagar por la persona con derecho a utilizar el terreno, que debe deducirlos de la renta que en otro caso estaría obligada a pagar al propietario.

Las *Rentas de terrenos* no incluyen los alquileres de los edificios y viviendas situados en ellos; tales alquileres se tratan como el pago de un servicio de mercado proporcionado por el propietario al arrendatario del edificio o de la vivienda y figuran en las cuentas como consumo intermedio o final de la unidad arrendataria. De no existir una base objetiva para dividir el pago entre rentas del terreno y alquiler de los edificios situados en él, se considera que el importe total es renta de terrenos cuando se estima que el valor del terreno es superior al de los edificios y que es alquiler de los edificios en caso contrario.

Las *Rentas de activos del subsuelo* comprenden los cánones y regalías que perciben los propietarios de los yacimientos de minerales y combustibles fósiles (carbón, petróleo o gas natural) a cambio de otorgar a otras unidades institucionales el derecho a explorar o explotar dichos yacimientos durante un período de tiempo especificado.

Por su parte, en la ECV, la variable asimilable a la rúbrica D.45, se recoge a través de la *Renta procedente del alquiler de una propiedad o terreno*, HY040G, referida a los ingresos percibidos en concepto de alquiler de una propiedad o terreno una vez deducidos los gastos, como el reembolso de los intereses de la hipoteca, las reparaciones menores, el mantenimiento, el seguro, etc³⁴. Pero incluyen, por ejemplo, tanto el alquiler de una vivienda no incluida en los beneficios y pérdidas de empresas no constituidas en sociedad, como los ingresos procedentes de inquilinos del alquiler de un terreno.

Por tanto, como podemos observar, la partida contable D.45 y la variable de la encuesta HY040G no recogen los mismos conceptos, ya que en la ECV se incluyen no sólo las rentas procedentes del alquiler de terrenos sino también las recibidas por el alquiler de viviendas por parte de los hogares, sean o no empresas no constituidas en sociedad, a precios de mercado (recordemos que si la vivienda ha sido alquilada a un precio inferior al de mercado o se encuentra ocupada a título gratuito, entonces se recoge en la ECV en la variable HY030N, *Alquiler imputado*).

En lo que respecta a la CNE, las rentas recibidas del alquiler de viviendas o partes de ellas pertenecientes a los hogares se registran como *Renta Mixta*. Las pagadas por este concepto se considerarían Gasto en consumo final de los hogares o Consumos intermedios para las empresas no constituidas en sociedad pertenecientes a los hogares. Para solucionar las diferencias anteriores y aproximarnos a las *Rentas de la tierra* de CNE, estimamos en la ECV los *Ingresos monetarios procedentes del alquiler de una propiedad o parte de ella*, IMALQPRO_B, como ya comentamos en el apartado dedicado a la *Renta Mixta*. De modo que la partida contable *Rentas de la tierra* (recursos menos empleos), la estimaremos en la ECV a través de la variable IMALQTER_B que, por tanto, será la diferencia entre HY040G y la variable estimada IMALQPRO_B.

Una vez estudiadas todas y cada una de las partidas de la Cuenta de asignación de la renta primaria para el sector hogares, nos faltaría para completarla el saldo de ésta, encargado de equilibrar el total de recursos y el total de empleos en dicha cuenta: el *Saldo de rentas primarias*. En CNE es igual a:

Saldo de rentas primarias Bruto (B.5b)

- = *Excedente Bruto de Explotación (B.2b_r)*
- + *Renta mixta bruta (B.3b_r) + Sueldos y Salarios (D.11_r)*
- + *Cotizaciones sociales a cargo de los empleadores (D.12_r)*
- + *Intereses recibidos (D.41_r)*
- + *Rentas distribuidas de las sociedades (D.42_r)*
- + *Otras rentas de inversión (D.44_r)*
- + *Rentas de la tierra recibidas (D.45_r)*
- *Intereses pagados (D.41_e)*
- *Rentas de la tierra pagadas (D.45_e)*

³⁴ Suponemos que entre dichos gastos están incluidos los pagos correspondientes por el derecho a la explotación de superficies de agua que CNE incluye en los empleos de la partida D.45, por lo que la variable HY040G de la ECV podría pensarse que se podría equiparar inicialmente a la diferencia entre los recursos y los empleos de las *Rentas de la tierra* de CNE.

donde ‘_r’ y ‘_e’ designa la posición de dicha partida contable en los recursos o empleos de dicha cuenta.

Dicho saldo, que designaremos en la ECV por medio de la variable B_{5b_Bdh} , se calcula como:

$$B_{5b_Bdh} = AIVPP_P - AIVPP_CI - AIVPP_OIN + CCFEE + RMCP_B \\ + RNMCP_B + CSEFCP + IMALQPRO_B + CCFRM + RMCA_B \\ + RNMCA_B + ORNC1_B + CSEFEMP + CSIMEMP + IMCAP_B \\ + INTP_B + ORI_B + IMALQTER_B$$

El Saldo de rentas primarias figura en el lado de los empleos de la Cuenta de asignación de la renta primaria y, en la cuenta siguiente, es decir, la Cuenta de distribución secundaria de la renta, sería la primera de las partidas (en el lado de los recursos) a raíz de la cual se construye esta última y que estudiaremos en el siguiente apartado.

En el cuadro 2.7 se presenta una síntesis sobre la construcción homogénea de cada una de las partidas de la Cuenta de asignación de la renta primaria que permiten obtener el *Saldo de rentas primarias Bruto* de CNE a partir de los microdatos de la ECV.

Cuadro 2.7 Cuenta de asignación de la renta primaria del sector hogares: correspondencia entre la CNE y la ECV

Concepto CNE	Código partida CNE (bruta)	Concepto ECV	Código variable ECV (bruta)
+ Excedente bruto de explotación (recursos)	B.2b_r	= EXCEDENTE BRUTO DE EXPLOTACIÓN	B.2b_Bdh
		+ Alquiler imputado asociado a la vivienda principal + Consumo de Capital Fijo asociado al Excedente de Explotación	(AIVPP_P - AIVPP_CLTCE) (=HY030N si HH020=1 en ECV09, y HY030N si HH021=1 ó 2 en ECV12) - HY100G - AIVPP_CLSeg* - AIVPP_OIN* CCFEE*
+ Renta mixta bruta (recursos)	B.3b_r	= RENTA MIXTA BRUTA	B.3b_Bdh
		+ Renta Bruta Monetaria de los trabajadores por cuenta propia + Renta Bruta No Monetaria de los trabajadores por cuenta propia + Cotizaciones sociales de los trabajadores por cuenta propia + Ingresos monetarios recibidos por el alquiler de propiedades (excluidos terrenos) + Consumo de Capital Fijo asociado a la Renta Mixta	RMCP_B = $\sum(PY050G) + \sum((\text{filtro1} \text{ ó } \text{filtro2}) \times PY010G)$ siendo, $\text{filtro1} = ((\text{situprop}=1 \text{ or } \text{situprop}=2) \text{ and } (PL100_F=-2) \text{ and } (PL073+PL074=0) \text{ and } (PY010G>0 \text{ or } PY020G>0))$; $\text{situprop}=1$ cuando $PL040_F=1$ and $PL040=1$ y $\text{situprop}=2$ si $PL040_F=1$ and $PL040=2$ $\text{filtro2} = ((\text{situprop}=1 \text{ or } \text{situprop}=2) \text{ and } (PL100_F=1 \text{ and } PL100>0) \text{ and } (PL073+PL074=0) \text{ and } (PY010G>0 \text{ or } PY020G>0))$ RNMCP_B = $\sum(PY070G)$ en ECV08 ó HY170N en ECV11 + $\sum((\text{filtro1} \text{ ó } \text{filtro2}) \times PY020G)$ CSEFCP = $\sum((\text{filtro1} \text{ ó } \text{filtro2}) \times PY030G)$ IMALQPRO_B* = HY040G - IMALQTER_B* CCFRM*
+ Remuneración de Asalariados (recursos)	D.1_r	= RENTAS SALARIALES	D.1_r_Bdh
+ Sueldos y Salarios (recursos)	D.11_r	= SUELDOS Y SALARIOS	D.11_r_Bdh
		+ Renta bruta monetaria o cuasi monetaria del asalariado + Renta bruta no monetaria del asalariado + Renta bruta del trabajo percibida por los menores de 16 años	RMCA_B = $\sum(PY010G) - \sum((\text{filtro1} \text{ ó } \text{filtro2}) \times PY010G)$ RNMCA_B = $\sum(PY020G) - \sum((\text{filtro1} \text{ ó } \text{filtro2}) \times PY020G)$ ORNC1_B*
+ Cotizaciones sociales a cargo de los empleadores (recursos)	D.12_r	= COTIZACIONES	D.12_r_Bdh
+ Cotizaciones sociales efectivas a cargo de los empleadores (recursos)	D.121_r**	+ Cotizaciones sociales efectivas a cargo de los empleadores	CSEFEMP = $\sum(PY030G) - CSEFCP$
+ Cotizaciones sociales imputadas a cargo de los empleadores (recursos)	D.122_r**	+ Cotizaciones sociales imputadas a cargo de los empleadores	CSIMEMP*
+ Rentas de la propiedad (recursos-empleos)	D.4_re	= RENTAS DE LA PROPIEDAD	D.4_re_Bdh
+ Intereses (recursos)	D.4142_r	+ Intereses, dividendos y ganancias brutos de inversiones de capital en empresas no constituidas en sociedad	IMCAP_B = HY090G
+ Rentas distribuidas de las sociedades (recursos)			
- Intereses (empleos)	D.41_e	- Intereses pagados	INTP_B*
+ Otras rentas de inversión (recursos)	D.44_r	+ Otras rentas de inversión recibidas	ORI_B*
+ Rentas de la tierra (recursos-empleos)	D.45_re	+ Rentas de la tierra (recibidas - pagadas)	IMALQTER_B* = HY040G-IMALQPRO_B*
= Saldo de rentas primarias bruto	B.5b	= SALDO DE RENTAS PRIMARIAS BRUTO	B.5b_Bdh

(*) : Variables no disponibles en la ECV pendientes de estimar en el desarrollo de este trabajo.

(**): Variables estimadas en CNE (base 2010) debido a las nuevas normas de publicación.

Fuente: Elaboración propia a partir de los datos del INE de la CNE (base 2010) y de las ECV.

3.1.2. Análisis de la homogeneidad entre la ECV y CNE relativa a las variables de ingresos: Cuenta de distribución secundaria de la renta

A continuación, nos centraremos en el análisis comparativo de las rúbricas que constituyen la Cuenta de distribución secundaria de la renta tanto en la CNE como en la ECV: *Saldo de rentas primarias, Impuestos corrientes sobre la renta, el patrimonio, etc., Cotizaciones sociales, Prestaciones sociales distintas de las transferencias sociales en especie y Otras transferencias corrientes.*

SALDO DE RENTAS PRIMARIAS

Como acabamos de ver, la primera componente de la cuenta de distribución secundaria de la renta se corresponde con el saldo de la cuenta de asignación de la renta primaria, esto es, el total de recursos menos el total de empleos, correspondiendo éste a la rúbrica B.5b de CNE y estimado en la ECV a través de la variable B.5b_Bdh.

Tras la incorporación del Saldo de Rentas Primarias (en el lado de los recursos), la Cuenta de distribución secundaria de la renta se completa con las partidas que se explican a continuación.

IMPUESTOS CORRIENTES SOBRE LA RENTA, EL PATRIMONIO, ETC.

Según el SEC 2010, los *Impuestos corrientes sobre la renta, el patrimonio, etc.* (D.5) comprenden todos los pagos obligatorios sin contrapartida, en efectivo o en especie, recaudados periódicamente por las administraciones públicas y por el resto del mundo sobre la renta y el patrimonio de las unidades institucionales, así como algunos impuestos periódicos que no se exigen ni sobre la renta ni sobre el patrimonio.

Los Impuestos corrientes sobre la renta, el patrimonio, etc., se dividen en Impuestos sobre la renta (D.51) y en Otros impuestos corrientes (D.59).

Los *Impuestos sobre la renta* (D.51) comprenden los impuestos sobre las rentas, los beneficios y las ganancias de capital. Se exigen sobre las rentas efectivas o presuntas de las personas físicas, los hogares, las sociedades y las instituciones sin fines de lucro. Incluyen los impuestos sobre la propiedad de bienes raíces e inmuebles cuando éstos se usan como base para estimar la renta de sus propietarios.

En concreto, la partida D.51 comprende:

- a) los impuestos sobre la renta individual o la renta del hogar (rentas del trabajo, de la propiedad, de la empresa, pensiones, etc.), incluidos los impuestos deducidos por los empleadores (impuestos retenidos en el momento de percibirse la renta) y los impuestos sobre la renta de los propietarios de empresas no constituidas en sociedad;
- b) los impuestos sobre la renta o los beneficios de las sociedades;
- c) los impuestos sobre las ganancias de posesión;
- d) los impuestos sobre los premios de loterías, juegos y apuestas, a pagar sobre las sumas recibidas por los ganadores, a diferencia de los impuestos sobre el volumen de negocio de los organizadores de los juegos de azar o loterías, que se tratan como impuestos sobre los productos.

En *Otros impuestos corrientes* (D.59) el SEC 2010 incluye:

- a) los impuestos corrientes sobre el capital, que consisten en impuestos a pagar sobre la propiedad o el uso de terrenos o edificios por sus propietarios, y los impuestos corrientes sobre el patrimonio neto y sobre otros activos (por ejemplo, joyas), excepto los impuestos mencionados en la rúbrica D.29 (que pagan las empresas por dedicarse a la producción) y en D.51 (impuestos sobre la renta);
- b) los impuestos de capitación que se exigen por persona adulta u hogar, independientemente de la renta o el patrimonio;
- c) los impuestos sobre el gasto, basados en el gasto total de la persona física o el hogar;
- d) los pagos de los hogares por licencias de tenencia y disfrute de vehículos, barcos o aviones, para usos no productivos, o de permisos de caza, pesca, etc.;
- e) los impuestos sobre operaciones internacionales, por ejemplo, viajes y remesas de fondos al extranjero, inversiones extranjeras, etc., excepto los que deben pagar los productores y los derechos de importación a cargo de los hogares.

Por su parte, los *Impuestos corrientes sobre la renta, el patrimonio, etc.*, no comprenden en ningún caso:

- los impuestos sobre las herencias, los derechos de sucesión y los derechos sobre donaciones inter vivos, que se considera que gravan el capital de los beneficios y que figuran en la rúbrica *Impuestos sobre el capital* (D.91);
- las exacciones ocasionales o excepcionales sobre el capital o el patrimonio, que figuran también en la partida D.91;
- los impuestos sobre los terrenos, edificios y otros activos propiedad de las empresas o que éstas alquilan y que son utilizados en sus procesos productivos, que se consideran *Otros impuestos sobre la producción* (D.29);
- los pagos de los hogares por licencias distintas de las de tenencia y disfrute de vehículos, barcos o aviones, o de los permisos de caza o pesca, es decir, permisos de conducir y de pilotar, cánones de radio y televisión, permisos de armas de fuego, entradas a museos y bibliotecas, tasas de recogida de basuras, etc. que en la mayoría de los casos se tratan como compras de servicios prestados por las administraciones públicas.

El importe total de los impuestos (D.5) incluye los intereses por mora y las multas impuestas por las autoridades fiscales, cuando no existen datos que permitan calcular por separado dichos intereses y multas. Asimismo, comprende las cargas impuestas en relación con la recaudación y liquidación de impuestos pendientes de pago, menos el importe de las desgravaciones fiscales concedidas por las administraciones públicas en el marco de su política económica y las devoluciones de impuestos por percepción indebida de los mismos.

En lo que respecta a la ECV, no podemos saber con exactitud cuál es el montante debido a los *Impuestos sobre la renta* equiparable al de CNE ya que se recogen junto a las *Cotizaciones sociales* efectivas por los hogares en la variable *Impuesto sobre la renta y cotizaciones sociales*, HY140G.

Los *Impuestos sobre la renta* se refieren en la ECV a aquéllos relativos a la renta, los beneficios y las ganancias de capital. Se establecen a partir de los ingresos reales o estimados de las personas, los hogares o la unidad familiar a efectos fiscales. Incluyen los impuestos estimados sobre propiedades, terrenos o bienes inmuebles cuando éstos sirven de base para el cálculo de la renta de sus propietarios.

En concreto, según la ECV, los impuestos sobre la renta incluyen:

- Los impuestos sobre la renta de las personas físicas (ingresos procedentes del empleo, la propiedad, la actividad de empresa, las pensiones, etc.), que incluyen los impuestos retenidos en origen por los empleadores, otras retenciones en origen y los impuestos sobre la renta de propietarios de empresas no constituidas en sociedad pagados durante el período de referencia de los ingresos.
- Las devoluciones/ingresos complementarios por ajustes en impuestos sobre la renta: consisten en el pago o la devolución de impuestos durante el período de referencia de los ingresos relativa a los impuestos abonados en concepto de la renta percibida durante el período de referencia de los ingresos o de la renta percibida en años anteriores.
- Los intereses de demora en el pago de impuestos y todas las multas impuestas por las autoridades fiscales.

Para poder comparar los *Impuestos sobre la renta* (IRPF) de CNE recogidos en la encuesta en HY140G junto a las cotizaciones sociales efectivas por los hogares, con los de CNE (D.51), deberemos estimar en la ECV dichas cotizaciones, para lo cual utilizaremos la variable CSEFHOG. Dichas cotizaciones se refieren a las pagadas por los asalariados, y por los trabajadores autónomos y los desempleados, que en CNE se contabilizan en las partidas D.6112 y D.6113, respectivamente. De esta forma, en la ECV tendremos que: $IRPF = HY140G - CSEFHOG$.

Respecto a los *Otros impuestos corrientes* (D.59) de CNE, la ECV proporciona la variable *Impuesto sobre el patrimonio*, HY120G, en la que incluye no sólo el impuesto sobre el patrimonio sino también otros impuestos corrientes. En concreto, la variable *Impuesto sobre el patrimonio* de la encuesta recoge, durante el período de referencia de los ingresos, el conjunto de impuestos que deben pagar periódicamente los propietarios en concepto de propiedad o utilización de terrenos o edificios, así como los impuestos corrientes sobre el patrimonio neto y otros activos (joyas u otros signos externos de riqueza). Además, al igual que D.59, HY120G incluye todos los intereses de demora en el pago del impuesto sobre el patrimonio y todas las multas impuestas por las autoridades fiscales pagados durante el período de referencia de los ingresos, así como los impuestos sobre la propiedad pagados directamente a la autoridad fiscal por el propietario durante el período de referencia de los ingresos.

No obstante, la ECV excluye de la variable HY120G:

- Los impuestos ocasionales, como los impuestos y derechos de sucesión o los derechos sobre donaciones.
- Los impuestos estimados sobre propiedades, terrenos o bienes inmuebles cuando éstos sirven de base para el cálculo de la renta de sus propietarios (estos impuestos se incluyen en la ECV en *Impuestos sobre la renta y cotizaciones sociales*).

- Los impuestos sobre terrenos, edificios y otros activos que las empresas tienen en propiedad o alquiler y que utilizan para la producción (estos impuestos se consideran impuestos sobre la producción y se deducen de la producción comercial de la renta de los trabajadores por cuenta propia para establecer la componente *Beneficios o pérdidas monetarios brutos de trabajadores por cuenta propia (incluidos derechos de propiedad intelectual)*)).

En definitiva, a la luz de las definiciones expuestas en las fuentes de estudio, podemos concluir que para elaborar con la ECV la partida contable D.5, *Impuestos corrientes sobre la renta, el patrimonio, etc.*, utilizaremos las variables de la ECV *Impuesto sobre la renta y cotizaciones sociales*, HY140G, menos una estimación de las *Cotizaciones Sociales Efectivas por los hogares* (CSEFHOG) y más HY120G, *Impuestos sobre el patrimonio*.

COTIZACIONES SOCIALES Y PRESTACIONES SOCIALES

En este apartado nos detendremos en el estudio de las Cotizaciones y Prestaciones Sociales. Ambas, se estudian de forma conjunta en el SEC 2010, rúbrica D.6, ya que no se puede entender el significado de las unas sin las otras, como a continuación veremos.

El SEC 2010 define las *Cotizaciones sociales netas* (D.61) como el conjunto de pagos que realizan los hogares a los sistemas de seguros sociales con el fin de asegurar el pago de prestaciones sociales.

Por su parte, las *Prestaciones sociales* (D.62) son transferencias a los hogares, en efectivo o en especie, destinadas a aligerar la carga financiera que representa para ellos la cobertura de una serie de riesgos o necesidades y efectuadas por medio de sistemas organizados colectivamente o, fuera de estos sistemas, por unidades de las administraciones públicas y por las ISFLSH. Dichas transferencias incluyen los pagos de las Administraciones Públicas a los productores que suministran prestaciones de forma individual a los hogares en el marco de los riesgos y necesidades sociales.

La relación de riesgos y necesidades que pueden dar origen a *Prestaciones sociales* en CNE es la siguiente:

- a) enfermedad,
- b) invalidez, incapacidad,
- c) accidente de trabajo o enfermedad profesional,
- d) vejez,
- e) supervivencia,
- f) maternidad,
- g) cargas familiares,
- h) fomento del empleo,
- i) desempleo,
- j) vivienda (alojamiento)³⁵,
- k) educación,
- l) pobreza.

³⁵ En el caso de la vivienda, los pagos efectuados por las administraciones públicas a los arrendatarios para reducir sus alquileres constituyen prestaciones sociales, excepto las prestaciones especiales pagadas por las administraciones públicas en su calidad de empleadores.

Asimismo, el SEC 2010 define los *sistemas de seguros sociales* como aquellos en los que los participantes se ven obligados a inscribirse o son alentados a hacerlo, por sus empleadores o por las administraciones públicas, con el fin de asegurarse contra determinados riesgos sociales o circunstancias que pueden perjudicar su bienestar o el de las personas a su cargo. En dichos sistemas, los asalariados u otros, o bien los empleadores en nombre de sus asalariados, pagan cotizaciones sociales para garantizar el derecho de los asalariados u otros cotizantes, y el de las personas a su cargo o las supérstites, a recibir prestaciones de seguros sociales en el ejercicio corriente o en otros posteriores.

Los sistemas de seguros sociales se organizan por grupos de trabajadores o están disponibles por ley para todos los trabajadores o categorías específicas de trabajadores, que incluyen tanto a los desempleados como a los asalariados. Pueden ser sistemas privados para grupos específicos de trabajadores empleados por un único empleador o sistemas de seguridad social que cubren al conjunto de la mano de obra de un país. La participación en dichos sistemas puede ser voluntaria para los trabajadores en cuestión, pero lo habitual es que sea obligatoria. Por ejemplo, la participación en sistemas organizados por empleadores particulares puede ser obligatoria con arreglo a las condiciones de empleo fijadas por un convenio colectivo entre los empleadores y sus asalariados.

Según CNE, los sistemas de seguros sociales se clasifican en dos tipos:

a) Sistemas de seguridad social, que cubren a toda la comunidad, o a una gran parte de ella, y que son impuestos, controlados y financiados por unidades de las administraciones públicas. Las pensiones a pagar con arreglo a dichos sistemas pueden guardar, o no, una relación con los niveles salariales del beneficiario o con su vida laboral. Las prestaciones distintas de las pensiones están con menos frecuencia vinculadas a los niveles salariales.

b) Sistemas vinculados al empleo, que derivan de una relación entre el empleador y el trabajador asalariado a la hora de conceder derechos por pensiones, y posiblemente de otro tipo, que forman parte de las condiciones de empleo y en los que la responsabilidad de la concesión de las prestaciones no recae en las administraciones públicas con arreglo a las normas en materia de seguridad social. Este sistema de seguros sociales es el que utilizan las unidades de las administraciones públicas para sus propios asalariados.

Cotizaciones sociales netas

El SEC 2010 define las cotizaciones sociales netas (D.61) como las *Cotizaciones sociales efectivas a cargo de los empleadores* (D.611), más las *Cotizaciones sociales imputadas a cargo de los empleadores* (D.612), más las *Cotizaciones sociales efectivas a cargo de los hogares* (D.613), más las *Cotizaciones sociales suplementarias a cargo de los hogares* (D.614), menos los *Servicios de los sistemas de seguros sociales* (D.61SC).

Las *Cotizaciones sociales efectivas a cargo de los empleadores* (D.611) corresponden al flujo D.121. que ya definimos páginas atrás como componente de la *Remuneración de asalariados*.

Los pagos de las cotizaciones sociales efectivas pueden ser obligatorios en virtud de una ley o un reglamento, o bien en virtud de convenios colectivos establecidos en determinadas ramas de actividad, de acuerdos entre el

empleador y los asalariados de una empresa concreta, o bien porque así consta en el propio contrato de trabajo. A veces, las cotizaciones pueden ser voluntarias, en cuyo caso comprenden:

- las cotizaciones sociales que las personas sin obligación legal de cotizar pagan a un fondo de seguridad social;
- las cotizaciones sociales pagadas a compañías de seguros (o a fondos de pensiones clasificados en el mismo sector) como parte de sistemas complementarios de seguro organizados por las empresas en beneficio de sus asalariados y a los que estos se adhieren voluntariamente;
- las cotizaciones pagadas a mutualidades a las que pueden adherirse los asalariados y los trabajadores por cuenta propia.

Las *Cotizaciones sociales imputadas a cargo de los empleadores* (D.612) representan la contrapartida de las prestaciones sociales directas de los empleadores (menos, en su caso, las cotizaciones sociales a cargo de los asalariados) pagadas directamente por los empleadores (es decir, no están vinculadas a las cotizaciones efectivas a cargo de los empleadores) a sus asalariados, antiguos asalariados y otros derechohabientes. Corresponden al flujo D.122 -descrito en el apartado dedicado a la *Remuneración de asalariados*- y su valor debe calcularse, en principio, mediante consideraciones actuariales o en un porcentaje razonable de los sueldos y salarios pagados a los asalariados actuales, o equivalente a las prestaciones distintas de las pensiones sin constitución de reservas que las empresas deben pagar durante el mismo ejercicio contable.

Las *Cotizaciones sociales efectivas a cargo de los hogares* (D.613) son cotizaciones sociales que pagan, en su nombre, los asalariados, los trabajadores por cuenta propia o los desempleados a los sistemas de seguros sociales.

Las *Cotizaciones sociales suplementarias a cargo de los hogares* (D.614) consisten en las rentas de propiedad obtenidas durante el ejercicio contable en el stock de los derechos por pensiones y otras prestaciones distintas de las pensiones. Este tipo de cotizaciones se incluyen en la renta de la propiedad que pagan los administradores de los fondos de pensiones a los hogares en la cuenta de asignación de la renta primaria (*Rentas de inversión a pagar sobre los derechos por pensiones*, D.442). Puesto que, en la práctica, los administradores de los fondos de pensiones retienen esta renta, en la cuenta de distribución secundaria de la renta se trata como un reembolso por los hogares a los fondos de pensiones en forma de cotizaciones sociales suplementarias a cargo de los hogares.

Por último, los *Servicios de los sistemas de seguros sociales* (D.61SC) son los costes del servicio cobrados por las unidades que administran los sistemas. Aunque figuran como parte del cálculo de las cotizaciones sociales netas (D.61), no son operaciones de redistribución, sino parte de los gastos de producción y consumo relacionados con estos servicios.

Una vez descritas las cotizaciones sociales en CNE, pasamos a analizar cómo se estimarían con los microdatos de la ECV:

- Según su metodología, la ECV nos proporciona el conjunto de cotizaciones sociales efectivas dentro de la variable PY030G, *Cotizaciones sociales a cargo del empleador*, las cuales, una vez deducidas las

abonadas por los trabajadores por cuenta propia, CSEFCP -que en la encuesta se incluyen también en esta variable-, darían como resultado la variable CSEFEMP que nos serviría para aproximarnos a la partida contable D.611.

- La *Cotizaciones sociales imputadas a cargo de los empleadores*, D.612, ya sabemos del apartado de la *Remuneración de asalariados* que no se recogen en la ECV, así que, como entonces propusimos, utilizaremos una estimación de las mismas por medio de la variable CSIMEMP.
- Las *Cotizaciones sociales efectivas a cargo de los hogares* de CNE pueden estimarse en la ECV como suma de las CSEFCP, incluidas en PY030G, y de las cotizaciones sociales efectivas a cargo de los asalariados y de los trabajadores autónomos y desempleados, CSEFHOG, que ya explicamos cuando estudiamos los *Impuestos sobre la renta* que se recogen en la variable *Impuesto sobre la renta y cotizaciones sociales*, HY140G.
- Las *Cotizaciones sociales suplementarias a cargo de los hogares* no se recogen en la ECV y, por tanto, tendremos que estimarlas. Para ello utilizaremos la variable CSSHOG.
- Finalmente, los *Servicios de los sistemas de seguros sociales* de CNE, D.61SC, de los que tampoco nos ofrece información la ECV, los estimaremos a través de la variable OCS -Otras cotizaciones sociales-.

Prestaciones sociales

En relación al conjunto de prestaciones sociales que reciben los hogares, Contabilidad Nacional, distingue entre las *Prestaciones sociales distintas de las transferencias sociales en especie* (D.62) y las *Transferencias sociales en especie* (D.63).

Las *Prestaciones sociales distintas de las transferencias sociales en especie* (D.62) son transferencias a los hogares, en efectivo o en especie, destinadas a aligerar la carga financiera que representa para ellos la cobertura de una serie de riesgos o necesidades y efectuadas por medio de sistemas organizados colectivamente o, fuera de estos sistemas, por unidades de las administraciones públicas y por las ISFLSH.

Por su parte, las *Transferencias sociales en especie* (D.63) comprenden los bienes y servicios individuales proporcionados a los hogares individuales, de forma gratuita o a precios económicamente no significativos, por unidades de las administraciones públicas y las ISFLSH. Este último tipo de prestaciones no se incluyen en la Cuenta de distribución secundaria de la renta ni se recogen en la ECV, sino que se encuadran en el proceso de redistribución de la misma, motivo por el cual no seguiremos profundizando en su estudio.

En CNE, las *Prestaciones sociales distintas de las transferencias sociales en especie* se clasifican en:

- a) *Prestaciones de seguridad social en efectivo* (D.621): son prestaciones de seguros sociales a pagar en efectivo a los hogares por los fondos de la seguridad social. Los reembolsos están excluidos y se tratan como transferencias sociales en especie (D.632). Dichas prestaciones se proporcionan en el marco de los sistemas de seguridad social y pueden dividirse en *Pensiones de la seguridad social en efectivo* (D.6211), y *Prestaciones distintas de las pensiones de la seguridad social en efectivo* (D.6212).

b) *Prestaciones de otros sistemas de seguros sociales (D.622)*: se corresponden con las prestaciones a pagar por los empleadores en el contexto de otros sistemas de seguros sociales vinculados al empleo. Los otros sistemas de seguros sociales vinculados al empleo son prestaciones sociales (en efectivo o en especie) a pagar por los sistemas de seguros sociales distintos de la seguridad social a los cotizantes al régimen, así como a las personas a su cargo y a sus supérstites. Entre ellas figuran, normalmente:

- los salarios normales o reducidos que se continúan pagando durante los períodos de ausencia del trabajo por enfermedad, accidente, maternidad, etc.;
- los complementos familiares, las ayudas a la enseñanza y otras en relación con las personas a cargo;
- las pensiones de jubilación o de supervivencia pagadas a los antiguos asalariados o a sus supérstites y las indemnizaciones por despido, invalidez, muerte por accidente, etc., pagadas a los asalariados o a sus supérstites (cuando estas prestaciones están vinculadas a convenios colectivos);
- los servicios médicos generales no relacionados con el trabajo del asalariado;
- las casas de convalecencia y las residencias de ancianos.

Las *Prestaciones de otros sistemas de seguros sociales* se dividen a su vez en *Pensiones de otros sistemas de seguros sociales (D.6221)* y *Prestaciones distintas de las pensiones, de otros sistemas de seguros sociales (D.6221)*.

c) *Prestaciones de asistencia social en efectivo (D.623)*: son transferencias corrientes pagadas por unidades de las administraciones públicas o de las ISFLSH a los hogares, con el fin de cubrir las mismas necesidades que las prestaciones de los seguros sociales, pero que no se enmarcan en un sistema de seguros sociales que requiera la participación normalmente a través de cotizaciones sociales. Por lo tanto, excluyen todas las prestaciones abonadas por fondos de la seguridad social. Las prestaciones de asistencia social pueden proporcionarse en las siguientes circunstancias:

- cuando no existe otro sistema de seguros sociales que cubra las circunstancias en cuestión;
- aunque existan uno o varios sistemas de seguros sociales, los hogares en cuestión no participan en ellos y no tienen derecho a las prestaciones;
- si las prestaciones de seguros sociales se consideran insuficientes para cubrir las necesidades en cuestión y las prestaciones de asistencia social se pagan como complemento;
- en el marco de la política social general.

Este tipo de prestaciones no incluye las transferencias corrientes pagadas como consecuencia de contingencias o circunstancias que normalmente no están cubiertos por los sistemas de seguros sociales (por ejemplo, las transferencias efectuadas en caso de catástrofes naturales, que se registran como *Otras transferencias corrientes, D.7*, o como *Otras transferencias de capital, D.99*).

Una vez visto el detalle de las Prestaciones sociales de CNE asimilables a las de la ECV, pasamos ver cómo se recogen en dicha encuesta.

El propio concepto de Prestaciones sociales en la encuesta es muy similar al de la rúbrica D.62 de CNE. Así, en la ECV se definen como las transferencias corrientes percibidas por los hogares durante el período de referencia de los

ingresos con objeto de aligerarlos de la carga financiera que suponen determinados riesgos o necesidades, realizadas a través de sistemas organizados colectivamente o por entidades estatales e instituciones sin fines de lucro al servicio de los hogares (ISFLSH).

De igual forma, como en el SEC, para que una transferencia sea considerada prestación social en la ECV, deberá cumplir uno de los dos requisitos siguientes:

- que su cobertura sea obligatoria (por Ley, reglamento o acuerdo en convenio colectivo) para el grupo en cuestión;
- y que esté basada en el principio de solidaridad social (es decir, si se trata de una pensión basada en el sistema de seguro, la prima y los derechos no son proporcionales a la exposición al riesgo individual de las personas cubiertas).

Las Prestaciones sociales en la ECV se desglosan en base al tipo de riesgo o necesidad que están destinadas a cubrir:

- a) Ayudas por familia/hijos.
- b) Ayudas para vivienda.
- c) Prestaciones por desempleo.
- d) Prestaciones por vejez.
- e) Prestaciones por supervivencia.
- f) Prestaciones por enfermedad.
- g) Prestaciones por invalidez.
- h) Ayudas para estudios.
- i) Exclusión social no clasificada en otro apartado.

Aunque, como vemos, la lista de contingencias que contemplan las prestaciones de la ECV coincide con la de CNE, la clasificación de dichas prestaciones es diferente en ambas fuentes, ya que el SEC subdivide las prestaciones atendiendo al organismo encargado de efectuar el desembolso de las mismas.

Además, a diferencia de CNE, en la ECV, aunque sí se recogen, no se consideran prestaciones sociales aquellas prestaciones procedentes de los sistemas en los que el beneficiario ha cotizado únicamente de manera voluntaria, independientemente de su empleador o de la Administración (denominadas *Pensiones periódicas de planes privados individuales [distintos de los incluidos en SEEPROS]*, variable PY080G).

A continuación, estudiaremos de forma detallada cada una de las prestaciones tal y como se recogen en la ECV.

AYUDAS POR FAMILIA/HIJOS. Se refieren a aquellas prestaciones que proporcionan ayuda económica a los hogares para criar a los hijos y a las personas que tienen a su cargo familiares distintos de los hijos. En concreto, este tipo de prestaciones incluye:

- La prestación de mantenimiento de renta en caso de parto: cantidad fija o en función de los ingresos, cuyo objeto es compensar al progenitor por la pérdida de ingresos debido a la ausencia del trabajo como consecuencia del nacimiento durante el período anterior y/o posterior al parto o como consecuencia de la adopción.

- La asignación por nacimiento: normalmente se concede en forma de cantidad a tanto alzado o de pagos fraccionados en caso de nacimiento o adopción.
- Las prestaciones del permiso de maternidad/paternidad: se concede a la madre o al padre en caso de interrupción del trabajo o de reducción de la jornada laboral con objeto de criar a un hijo, normalmente de corta edad.
- El subsidio familiar o la prestación por hijo: pagos periódicos a un miembro del hogar con hijos a cargo para ayudarle a costear los gastos de educación de los hijos.
- Otras prestaciones monetarias: prestaciones percibidas con independencia de los subsidios familiares, con objeto de ayudar a los hogares a hacer frente a determinados gastos, como los derivados de las necesidades específicas de las familias monoparentales, familias con hijos discapacitados o personas que tienen a su cargo familiares distintos de los hijos. Dichas prestaciones pueden percibirse periódicamente o en forma de cantidad a tanto alzado.

Las *Prestaciones para familia/hijos* se recogen en la ECV a través de la variable HY050G. A lo largo de este trabajo las designaremos como PSFAM_B.

AYUDAS PARA VIVIENDA. Las *Ayudas para vivienda*, recogidas en la ECV por medio de la variable HY070G, se refieren a la intervención de las autoridades públicas para ayudar a los hogares a hacer frente a los gastos asociados a la vivienda. Un criterio esencial para definir el ámbito de aplicación del subsidio de vivienda es la existencia de unos requisitos de recursos económicos que deben cumplirse para determinar la asignación de la prestación. Incluyen:

- Subvención de alquiler: transferencia corriente sometida a unos requisitos de recursos económicos que concede una autoridad pública a un inquilino, temporalmente o a largo plazo, para ayudarle a hacer frente a los gastos de alquiler.
- Subvención a propietarios que residen en sus viviendas: transferencia sometida a unos requisitos de recursos económicos que concede una autoridad pública a un propietario que reside en su vivienda para aliviar los gastos corrientes relacionados con la misma (en la práctica se le suele ayudar con el pago de hipotecas y/o intereses).

Por el contrario, las *Ayudas para vivienda* de la ECV excluyen las políticas sociales de vivienda organizadas a través del sistema fiscal (es decir, las ventajas fiscales), y todas las *Transferencias de capital* (en particular, las *Ayudas a la inversión*).

En adelante, adoptaremos la nomenclatura PSVIV_B para referirnos a este tipo de prestaciones.

PRESTACIONES POR DESEMPLEO. Las *Prestaciones por desempleo* se refieren a las prestaciones que sustituyen, íntegra o parcialmente, la pérdida de renta de un trabajador debido a la pérdida de un empleo remunerado. Proporcionan una renta de subsistencia (o superior) a personas que se integran o reintegran en el mercado laboral; compensan por la pérdida de ingresos derivada del desempleo parcial; sustituyen, íntegra o parcialmente, la pérdida de renta de un trabajador que se jubila de un empleo remunerado antes de la edad legal de jubilación como consecuencia de una reducción de plantilla por motivos

económicos; contribuyen a costear los gastos de formación o reciclaje de personas que buscan trabajo; o ayudan a las personas desempleadas a hacer frente a los gastos de viaje o traslado para conseguir trabajo.

Incluyen:

- La prestación de desempleo total: prestación que compensa la pérdida de ingresos cuando una persona puede trabajar y está disponible pero no encuentra un empleo adecuado, incluidas las personas que no han trabajado anteriormente.
- La prestación de desempleo parcial: prestación que compensa la pérdida de sueldos o salarios como consecuencia de contratos laborales formales de jornada reducida y/o programas de trabajo intermitente, independientemente del motivo (recesión o desaceleración en la actividad de la empresa, avería del material, condiciones climáticas, accidentes, etc.), cuando se mantiene la relación empleador/asalariado.
- La jubilación anticipada por motivos del mercado laboral: pagos periódicos a los trabajadores mayores que se jubilan antes de alcanzar la edad normal de jubilación por motivos de desempleo o de reducción de plantilla provocados por medidas económicas como la reestructuración de un sector industrial o una empresa. Dichos pagos normalmente cesan cuando el beneficiario obtiene el derecho a una pensión de jubilación.
- Ayudas de formación profesional: pagos procedentes de los fondos de entidades públicas destinados a grupos de personas en la fuerza de trabajo que participan en programas de formación con objeto de desarrollar su potencial para el empleo.
- La movilidad y la reinstalación: pagos procedentes de los fondos de la seguridad social o de entidades públicas destinados a personas sin empleo con objeto de motivarlas a cambiar de localidad o de trabajo a fin de buscar o encontrar trabajo.
- Las indemnizaciones por cese: prestaciones que compensan a los asalariados por la interrupción de su trabajo antes de haber alcanzado la edad normal de jubilación para el trabajo desempeñado.
- La indemnización por despido: sumas de capital que se conceden a los asalariados que han sido despedidos por causa no imputable al trabajador, por parte de una empresa que suspende o reduce su actividad.
- Otras prestaciones monetarias: otras ayudas económicas, en particular el subsidio de desempleo de larga duración.

Por su parte, este tipo de prestaciones excluye las ayudas familiares adicionales concedidas por hijos a cargo (se incluyen en *Ayudas por familia/hijos*).

En la ECV las *Prestaciones por desempleo* se recogen por medio de la variable PY090G, observada para cada uno de los adultos del hogar. Para recoger la cantidad agregada de esta variable para cada hogar encuestado utilizaremos PSDES_B, siendo $PSDES_B = \sum (PY090G)$, suma esta extendida a todos los miembros adultos del hogar.

PRESTACIONES POR JUBILACIÓN. Este tipo de prestaciones se refiere en la ECV a la protección social que se concede para afrontar el riesgo vinculado a la edad, la pérdida de ingresos, la renta inadecuada, la falta de independencia para llevar a cabo las tareas cotidianas, la participación reducida en la vida

social, etc. Asimismo, las prestaciones por vejez incluyen aquellas que proporcionan una renta de sustitución cuando la persona mayor se retira del mercado laboral o garantizan una renta determinada cuando una persona ha alcanzado una determinada edad.

Además, las *Prestaciones por jubilación* incluyen:

- Pensiones de jubilación: pagos periódicos con objeto de mantener la renta del beneficiario tras su jubilación de un trabajo remunerado a la edad normal.
- Pensiones de jubilación anticipada: pagos periódicos con objeto de mantener la renta de los beneficiarios que se jubilan antes de alcanzar la edad normal tal y como se define en el plan correspondiente o en el plan de referencia. Puede concederse con o sin reducción de la pensión normal.
- Pensiones de jubilación parcial: pago periódico de un porcentaje de la pensión de jubilación total destinado a los trabajadores de más edad que continúan trabajando, pero reducen la jornada laboral o cuya renta procedente de una actividad profesional está por debajo de un límite establecido.
- Las asignaciones por cuidados: prestaciones concedidas a personas mayores que necesitan asistencia frecuente o constante para ayudarlas a hacer frente a los gastos extraordinarios de atención (distintos del cuidado médico) cuando la prestación no constituye el reembolso de un gasto certificado.
- Las prestaciones por supervivencia abonadas tras la edad normal de jubilación.
- Las prestaciones monetarias de invalidez abonadas tras la edad normal de jubilación.
- Los pagos a tanto alzado en la fecha normal de jubilación.
- Otras prestaciones monetarias: otras prestaciones periódicas y a tanto alzado abonadas en el momento de la jubilación o en concepto de vejez, como sumas de capital abonadas a personas que no cumplen todos los requisitos para recibir una pensión de jubilación periódica o que fueron miembros de un plan diseñado para proporcionar únicamente sumas de capital en el momento de la jubilación.

Y excluyen:

- Las ayudas familiares concedidas por hijos a cargo (se incluye en *Ayudas por familia/hijos*).
- Las prestaciones por jubilación anticipada por motivos del mercado laboral o en caso de disminución de la capacidad de trabajo (se incluyen, respectivamente, en «Prestaciones por desempleo» y en «Prestaciones por invalidez»).
- Las prestaciones concedidas a personas mayores que necesitan asistencia frecuente o constante para ayudarlas a hacer frente a los gastos extraordinarios de atención cuando la prestación constituye el reembolso de un gasto certificado.

Las *Prestaciones por jubilación* se recogen para cada adulto en la ECV en la variable PY100G. Para obtener esta variable agregada para todos los miembros de cada hogar encuestado utilizaremos la notación PSJUB_B, siendo $PSJUB_B = \sum (PY100G)$, suma esta aplicada a todos los miembros adultos del hogar.

PRESTACIONES POR SUPERVIVENCIA. Este tipo de prestaciones por supervivencia se refieren a aquéllas que aportan una renta temporal o permanente a personas que no han alcanzado la edad de jubilación y han sufrido la pérdida del cónyuge, la pareja o el pariente más cercano, generalmente cuando éste último representaba el sustento principal del beneficiario. Los supervivientes con derecho a la prestación pueden ser el cónyuge o ex cónyuge de la persona fallecida, sus hijos, sus nietos, sus padres y demás parientes. En algunos casos, la prestación también puede concederse a una persona ajena a la familia.

Normalmente, la *Prestación por supervivencia* se concede sobre la base de un derecho derivado, es decir, un derecho que en origen pertenecía a otra persona, cuyo fallecimiento es condición para que se conceda la prestación.

En concreto, las *Prestaciones por Supervivencia* incluyen:

- La pensión de supervivencia: pagos periódicos a personas cuyo derecho deriva de su relación con una persona fallecida cubierta por un esquema (viudos, viudas, huérfanos y similares).
- El subsidio de defunción: un único pago a una persona cuyo derecho deriva de su relación con una persona fallecida (viudos, viudas, huérfanos y similares).
- Otras prestaciones monetarias: otros pagos periódicos o cantidades a tanto alzado concedidos en virtud del derecho derivado de un superviviente.

Y excluyen:

- Las ayudas familiares concedidas por hijos a cargo (estas prestaciones se incluyen en «Ayudas por familia/hijos»).
- Gastos de sepelio.
- Prestaciones de supervivencia abonadas después de la edad normal de jubilación (estas prestaciones se incluyen en Prestaciones por vejez).

Las *Prestaciones por Supervivencia* se recogen para cada adulto en la ECV en la variable PY110G. Para agregarla para todos los adultos de cada hogar de la encuesta, utilizaremos la variable $PSSUP_B = \sum(PY110G)$, suma esta aplicada a todos los miembros adultos del hogar.

PRESTACIONES POR ENFERMEDAD. Las *Prestaciones por enfermedad* se refieren a las prestaciones monetarias que sustituyen, íntegra o parcialmente, la pérdida de ingresos durante un período de incapacidad temporal para el trabajo, por causa de enfermedad o lesión.

Las Prestaciones por enfermedad comprenden:

- La licencia por enfermedad: pagos de una cantidad fija o en función de los ingresos, cuyo objeto es compensar al beneficiario, íntegra o parcialmente, de la pérdida de ingresos motivada por su incapacidad temporal para el trabajo por causa de enfermedad o lesión. Dichas prestaciones pueden ser abonadas por sistemas de seguros sociales autónomos o por el empleador en forma de pago continuado de los sueldos y salarios durante el tiempo que dure la enfermedad.
- La licencia pagada en caso de enfermedad o lesión de un hijo a cargo.

- Otras prestaciones monetarias: diversos pagos destinados a las personas cubiertas en relación con enfermedad o lesión.

Se excluyen:

- Las prestaciones monetarias que sustituyen la pérdida de ingresos durante un período de incapacidad temporal para el trabajo en caso de embarazo (estas prestaciones se incluyen en «Ayudas por familia/hijos»).
- Las prestaciones monetarias que sustituyen la pérdida de ingresos durante un período de incapacidad temporal para el trabajo en caso de invalidez (estas prestaciones se incluyen en *Prestaciones por invalidez*).

En la ECV las Prestaciones por enfermedad se recogen para cada adulto mediante la variable PY120G. El total para el hogar de este tipo de prestaciones lo designaremos como $PSENF_B = \sum(PY120G)$, suma esta aplicada a todos los miembros adultos del hogar.

PRESTACIONES POR INVALIDEZ. Las prestaciones por invalidez se refieren a aquéllas que aportan una renta a personas que no han alcanzado la edad normal de jubilación y cuya capacidad para trabajar y percibir ingresos se ha visto deteriorada por encima de un nivel mínimo establecido por la Ley a causa de una incapacidad física o mental.

Se considera invalidez la incapacidad total o parcial para desarrollar una actividad económica o llevar una vida normal debido a un deterioro físico o mental que puede ser permanente o persistir por encima de un período mínimo determinado.

Incluyen:

- La pensión de invalidez: pago periódico cuyo objeto es mantener o completar la renta de una persona que no ha alcanzado la edad normal de jubilación y padece una invalidez que deteriora su capacidad para trabajar por encima de un nivel mínimo establecido por la Ley.
- La jubilación anticipada por reducción de la capacidad de trabajo: pagos periódicos a los trabajadores mayores que se jubilan antes de alcanzar la edad normal de jubilación como consecuencia de una reducción de la capacidad de trabajo. Dichos pagos normalmente cesan cuando el beneficiario obtiene el derecho a una pensión de jubilación.
- Ayuda asistencial: prestación que se concede a las personas inválidas que no han alcanzado la edad normal de jubilación y necesitan asistencia frecuente o constante con objeto de ayudarles a hacer frente a los gastos extras de asistencia (distinta de los cuidados médicos). Dicha prestación no debe constituir el reembolso de un gasto certificado.
- La integración económica de las personas con minusvalía: asignaciones que se conceden a personas discapacitadas cuando emprenden un trabajo adaptado a su situación, normalmente en un taller protegido, o siguen una formación profesional.
- Las prestaciones de invalidez para niños discapacitados a los que les corresponden por derecho propio, independientemente de su situación de dependencia.
- Otras prestaciones monetarias: pagos periódicos o cantidades a tanto alzado que no entran en los epígrafes anteriores, como ayudas económicas ocasionales, etc.

Por el contrario, las prestaciones por invalidez excluyen:

- Las prestaciones concedidas para sustituir, íntegra o parcialmente, los ingresos durante el período de incapacidad temporal para el trabajo por causa de enfermedad o lesión (estas prestaciones se incluyen en *Prestaciones por enfermedad*).
- Ayudas concedidas a las familias con algún miembro discapacitado (estas prestaciones se incluyen en *Ayudas por familia/hijos*).
- Las prestaciones concedidas a los supervivientes de personas discapacitadas, como las pensiones (estas prestaciones se incluyen en *Prestaciones por supervivencia*).
- Las prestaciones que constituyen el reembolso de gastos certificados.
- Las prestaciones de invalidez monetarias abonadas después de alcanzar la edad normal de jubilación (estas prestaciones se incluyen en *Prestaciones por vejez*).

En la ECV las Prestaciones por invalidez se recogen personalmente en la variable PY130G. Su agregado para todos los adultos de cada hogar representado en la encuesta lo obtendremos como $PSINV_B = \sum(PY120G)$, suma esta aplicada a todos los miembros adultos del hogar.

AYUDAS PARA ESTUDIOS. Las asignaciones por estudios se refieren a subvenciones, becas y demás ayudas para estudios que reciben los estudiantes. En la ECV las Ayudas para estudios se recogen a nivel de persona adulta en la variable PY140G y para obtener el total para cada hogar utilizaremos la variable $PSEST_B = \sum(PY140G)$, suma esta aplicada a todos los miembros adultos del hogar.

EXCLUSIÓN SOCIAL NO CLASIFICADA EN OTRO APARTADO. Las prestaciones sociales que se incluyen en la función *Exclusión social no clasificada en otro apartado* se refieren a los “excluidos sociales” o a “aquellos que corren el riesgo de convertirse en excluidos sociales”. El carácter general de este apartado permite identificar diversos grupos objetivo como indigentes, emigrantes, refugiados, toxicómanos, alcohólicos, víctimas de la violencia, entre otros.

Incluye:

- Ayuda económica: pagos periódicos destinados a personas con recursos insuficientes. Las condiciones para tener derecho pueden estar relacionadas no sólo con los recursos personales, sino también con la nacionalidad, la residencia, la edad, la disponibilidad para trabajar o la situación familiar. Estas prestaciones pueden tener una duración limitada o ilimitada; pueden concederse a una persona o a toda la familia y proceder de la Administración.
- Otras prestaciones monetarias: complemento para los indigentes y personas vulnerables con objeto de contribuir a paliar la pobreza o ayudar en las situaciones difíciles. Estas prestaciones pueden proceder de organizaciones privadas sin ánimo de lucro.

En la ECV este tipo de prestaciones se recogen de forma agregada para el hogar a través de la variable *Ingresos por asistencia social* (HY060G).

En definitiva, podemos ver que tanto esta encuesta como la CNE cubren un conjunto de riesgos o necesidades idénticos. No obstante, la diferencia entre ambas fuentes es debida a que en CNE los componentes de dichas prestaciones sociales se clasifican en base al tipo de sistema de seguro social que las

proporciona, mientras que en la ECV atiende a la clase de riesgo o necesidad que están destinadas a cubrir. Como consecuencia de ello, no va a ser posible la comparación de los distintos tipos de prestaciones en ambas fuentes, si no que tendremos que conformarnos con un análisis global.

En concreto, para hallar en la ECV el total de recursos de las *Prestaciones sociales distintas de las transferencias sociales en especie* relativas a las variables D.621 y D.623 de CNE (*Prestaciones de seguridad social en efectivo y de asistencia social en efectivo*, respectivamente):

- Agregaremos la serie de Prestaciones sociales recogidas como tal en la ECV: *Prestaciones por desempleo, Prestaciones por jubilación, Prestaciones por supervivencia, Prestaciones por enfermedad, Prestaciones por invalidez, Ayuda para estudios, Ayuda por familia/hijos, Ayuda para vivienda e Ingresos por asistencia social*.
- Habrá que eliminar, por otro lado, la parte de las Indemnizaciones de seguros no vida que los hogares reciben por daños personales, las cuales se incluyen en la ECV dentro de las *Prestaciones por invalidez o enfermedad*. Para estimarla utilizaremos la variable INDNOVIDA1_B. Este tipo de indemnizaciones se incluyen en CNE junto a las recibidas por los hogares en concepto de Indemnizaciones de seguros no vida por daños en los bienes en la partida contable D.72 *Indemnizaciones de seguro no vida*, que estudiaremos a continuación en el bloque de *Otras transferencias corrientes*.
- Y también deberemos incluir parte de las rentas que se recogen en la encuesta como 'Otras rentas no clasificadas' en la variable *Rentas percibidas por los menores de 16 años* (HY110G). A excepción de la parte destinada a las Rentas del trabajo (ORNC1_B), a la que ya nos referimos en la Remuneración de asalariados, la cuantía restante de HY110G suponemos que se atribuye a Prestaciones de Supervivencia (ORNC2_B), Prestaciones por enfermedad e invalidez (ORNC3_B) y Ayudas para estudios (ORNC4_B).

Por su parte, tal y como dice el SEC 2010, las *Cotizaciones sociales imputadas a cargo de los empleadores* (D.122) -estimadas en la ECV mediante CSIMEMP- representan la contrapartida de Prestaciones de otros sistemas de seguros sociales (D.622), por lo que dicha partida, que hemos designado en la encuesta como POSSS_r, tendrá el mismo importe que CSIMEMP.

El total de empleos de las *Prestaciones sociales distintos de las transferencias sociales en especie* de CNE (D.62_e), suponemos que se explica por las prestaciones que los hogares con trabajadores por cuenta propia prestan a los hogares. Desconocemos la cuantía de esta variable ya que la ECV sólo recoge las Prestaciones sociales recibidas por los hogares, por lo que para su incorporación en este trabajo procederemos a su estimación mediante la variable POSSS_e.

OTRAS TRANSFERENCIAS CORRIENTES

Las *Otras transferencias corrientes* (D.7) corresponden en el SEC 2010 a todas las transferencias corrientes distintas de las rentas de la propiedad, las pensiones y la renta de las personas que viven en instituciones de forma permanente. En el

caso particular del sector hogares, este tipo de transferencias se dividen en: *Primas netas de seguro no de vida* (D.71), *Indemnizaciones de seguro no de vida* (D.72) y *Transferencias corrientes diversas* (D.75).

Primas netas de seguro de vida

En Contabilidad Nacional las *Primas netas de seguro no de vida* (D.71) se definen como las primas a pagar en virtud de pólizas suscritas por unidades institucionales. Las pólizas de los hogares individuales son las que éstos suscriben por su propia iniciativa y en su propio beneficio, independientemente de los empleadores o de las administraciones públicas y fuera de todo sistema de seguros sociales. Las primas netas de seguro no de vida comprenden tanto las primas efectivas que deben pagar los tomadores de las pólizas para estar cubiertos durante el período contable (primas imputables al ejercicio corriente) como las primas complementarias a pagar de las rentas de la propiedad atribuidas a los tomadores del seguro, previa deducción de los costes del servicio prestado por las compañías de seguros.

Las *Primas netas de seguro no vida* son los montantes disponibles para la cobertura contra diversas contingencias o accidentes, de origen natural o causado por el hombre, que provocan daños a los bienes, a las propiedades o a las personas (incendios, inundaciones, accidentes, colisiones, naufragios, robos, actos violentos, enfermedades, etc.) o bien para la cobertura de las pérdidas financieras ocasionadas por contingencias, tales como enfermedades, desempleo, accidentes, etc.

En la ECV no se recogen las *Primas netas de seguro no vida*. Para estimarlas, recurriremos a otras fuentes de información y utilizaremos la variable PRIMNOVIDA_B.

Indemnizaciones de seguro no vida

Las *Indemnizaciones de seguro no de vida* (D.72), según el SEC 2010, representan las indemnizaciones imputables al ejercicio corriente en virtud de contratos de seguros no de vida; es decir, los importes que las compañías de seguro están obligadas a pagar por lesiones o daños sufridos por las personas o los bienes (incluidos los bienes de capital fijo).

Las indemnizaciones de seguro no de vida no incluyen los pagos que constituyen prestaciones sociales.

La liquidación de una indemnización de seguro no vida se considera una transferencia al demandante de la indemnización. Estos pagos se consideran transferencias corrientes, incluso cuando puede tratarse de importes elevados que resultan de la destrucción accidental de un activo fijo o de lesiones corporales graves a una persona. Las indemnizaciones por importes excepcionalmente elevados, como tras una catástrofe, pueden tratarse como *Transferencias de capital* en vez de *Transferencias corrientes*.

En general, las indemnizaciones recibidas por los demandantes no tienen que destinarse a ningún propósito en particular, y tampoco es necesario reparar o sustituir los bienes o activos dañados o destruidos.

Las indemnizaciones tienen su origen en los daños o lesiones que los tomadores del seguro causan a terceros o a las propiedades de estos. En tales

casos, se considera que las compañías de seguros pagan directamente las indemnizaciones a las partes perjudicadas y no indirectamente a través del tomador del seguro.

Tal y como hemos comentado con anterioridad, parte de las *Indemnizaciones de seguro no vida* recibidas por los hogares se recogen en la ECV en el apartado de *Prestaciones por enfermedad* y *Prestaciones por invalidez*, en concreto, las percibidas como consecuencia de daños personales (INDNOVIDA1_B). El resto de Indemnizaciones de seguros no vida, las recibidas por los hogares en virtud de los daños sufridos en sus bienes, no están recogidas en la ECV, por lo que procederemos a su estimación a través de INDNOVIDA2_B. Por ejemplo, en el caso de un accidente de tráfico, la parte recibida para la reparación del vehículo no se incluye en la encuesta. En cambio, las Indemnizaciones de seguros no vida por daños personales (por ejemplo, si la persona se rompe una pierna) se consideran prestaciones (o bien por enfermedad o por invalidez).

Transferencias corrientes diversas

En el SEC 2010 se distinguen diferentes tipos de Transferencias corrientes diversas relativas al sector hogares: *Transferencias corrientes a las ISFLSH* (D.751), *Transferencias corrientes entre los hogares* (D.752), y *Otras transferencias corrientes diversas* (D.759). A continuación, estudiamos cada una de ellas por separado.

Las *Transferencias corrientes a las ISFLSH* (D.751) comprenden todas las contribuciones voluntarias (excepto las herencias), las cuotas de sus miembros y la asistencia financiera que las ISFLSH reciben de los hogares (incluidos los hogares no residentes) y, en menor medida, de otras unidades.

En concreto, en la rúbrica D.751 se registran:

- las cuotas periódicas pagadas por los hogares a sindicatos y organizaciones políticas, deportivas, culturales, religiosas y similares clasificadas en el sector de las ISFLSH;
- las contribuciones voluntarias (excepto las herencias) de los hogares, las sociedades y el resto del mundo a las ISFLSH. Se incluyen aquí las transferencias en especie hechas a las organizaciones benéficas, en forma de donaciones de alimentos, ropa, mantas, medicinas, etc., para su distribución a los hogares residentes o no residentes. Dicho tratamiento se aplica a los bienes de consumo, ya que las transferencias de donaciones de gran cuantía (objetos valiosos tratados como activos no financieros) se registran en *Otras transferencias de capital* (D.99).
- las ayudas de las administraciones públicas, excepto las transferencias hechas específicamente para financiar gastos de capital, que forman parte de las ayudas a la inversión.

Se excluyen de las *Transferencias corrientes a las ISFLSH* los pagos de cuotas o suscripciones a instituciones sin fines de lucro al servicio de las empresas, tales como cámaras de comercio o asociaciones empresariales, que se consideran pagos por los servicios prestados.

Las *Transferencias corrientes entre los hogares* (D.752) consisten en todas las transferencias corrientes, en efectivo o en especie, que los hogares residentes efectúan a, o reciben de, otros hogares residentes o no residentes. En particular, incluyen las remesas de fondos que los emigrantes o los trabajadores instalados

permanentemente en el extranjero (o que trabajan en el extranjero durante un año o más) hacen a los miembros de su familia residentes en su país de origen, y las de los padres a sus hijos residentes en otros lugares.

Por su parte, las *Otras Transferencias corrientes* (D.759) comprenden:

- a) **Multas y sanciones.** Las multas y sanciones impuestas a las unidades institucionales por tribunales de justicia u otras instancias jurídicas se consideran *Transferencias corrientes diversas* (D.759). No se incluyen en esta rúbrica: las multas y sanciones impuestas por las autoridades fiscales por evasión o retraso en el pago de impuestos, que normalmente no pueden diferenciarse de los propios impuestos; ni las tarifas pagadas para la obtención de licencias, que se consideran o bien impuestos o bien pagos por los servicios prestados por unidades de las Administraciones Públicas.
- b) **Loterías y juegos de azar.** Los importes pagados por los billetes de lotería o las sumas colocadas en apuestas comprenden dos elementos: el pago del servicio a la unidad que organiza la lotería o los juegos de azar y una transferencia corriente residual que se paga a los ganadores. El coste del servicio puede ser muy importante y tiene que cubrir los impuestos sobre la producción de servicios de juegos de azar. En el sistema, se considera que las transferencias tienen lugar directamente entre quienes participan en la lotería o en los juegos de azar, es decir, entre los hogares. Cuando participan hogares no residentes, pueden darse transferencias netas significativas entre el sector hogares y el resto del mundo. Este tipo de transferencias se registran en el momento en que se efectúan.
- c) **Pagos compensatorios.** Los pagos compensatorios son las transferencias corrientes pagadas por unas unidades institucionales a otras en compensación por las lesiones o daños causados a las personas o a las propiedades, excluidas las indemnizaciones de seguro no de vida. Los pagos compensatorios pueden ser o bien pagos obligatorios dictados por un tribunal de justicia o bien pagos graciables acordados extrajudicialmente. Esta rúbrica comprende también los pagos voluntarios efectuados por unidades de las AA.PP. e ISFLSH en compensación por las lesiones o daños causados por catástrofes naturales, con excepción de los pagos clasificados como *Transferencias de capital*.
- d) **Otras transferencias corrientes.** Para los hogares se incluyen aquí:
 - o las becas, bolsas de viaje y recompensas pagadas a los hogares residentes y no residentes por las AA.PP. o las ISFLSH;
 - o las primas de ahorro concedidas regularmente por las AA.PP. a los hogares como recompensa por sus ahorros durante el período;
 - o las devoluciones hechas por los hogares de gastos efectuados en su favor por organizaciones de asistencia social;
 - o las transferencias corrientes de las AA.PP. a los hogares en su condición de consumidores, si no se registran como prestaciones sociales.

De las *Transferencias corrientes diversas* de CNE estudiadas, sólo se recogen un subconjunto de ellas en la ECV. En particular, se registran en las variables *Transferencias periódicas monetarias percibidas de otros hogares* (HY080N) y *Transferencias periódicas monetarias abonadas a otros hogares* (HY130N).

La ECV denomina *Transferencias periódicas monetarias percibidas de otros hogares* a las cantidades periódicas monetarias percibidas durante el período de referencia de los ingresos procedentes de otros hogares o personas. Incluyen:

- La pensión de manutención y las ayudas obligatorias a los hijos.
- La pensión de manutención y las ayudas voluntarias a los hijos, recibidas de manera periódica.
- La ayuda monetaria periódica procedente de personas distintas de los miembros del hogar.
- La ayuda monetaria periódica procedente de hogares residentes en otros países.

No obstante, las *Transferencias periódicas monetarias percibidas de otros hogares* excluyen la vivienda subvencionada o proporcionada de forma gratuita por otro hogar (se incluye en la variable *Alquiler imputado*).

Por su parte, las *Transferencias periódicas pagadas a otros hogares* se refieren en la encuesta a las cantidades periódicas monetarias abonadas a otros hogares durante el período de referencia de los ingresos. En concreto, la ECV recoge dentro de esta partida:

- La pensión de manutención y las ayudas obligatorias a los hijos.
- La pensión de manutención y las ayudas voluntarias a los hijos, abonadas de manera periódica.
- La ayuda monetaria periódica destinada a personas distintas de los miembros del hogar.
- La ayuda monetaria periódica destinada a hogares residentes en otros países.

En definitiva, como acabamos de ver, en la ECV sólo se registran una parte de las transferencias corrientes diversas de CNE: las que tienen lugar de forma periódica. De este modo, no tienen representación en la encuesta aquellas transferencias recibidas o pagadas puntualmente por los hogares, como las multas o sanciones, loterías y juegos de azar y los pagos compensatorios, entre otras. Para subsanar esta falta de información en la ECV, procederemos a su estimación por medio de las variables *Otras Transferencias recibidas por los hogares* (OTRDOH_B) y *Otras transferencias pagadas por los hogares* (OTPAOH_B).

Así pues, con las particularidades entre una y otra fuente de estudio que acabamos de mencionar en este último bloque, el conjunto de recursos de la partida *Otras transferencias corrientes* (D.7_r) de CNE la vamos a aproximar en nuestro estudio a través de la variable de la ECV *Transferencias periódicas monetarias percibidas de otros hogares*, HY080N, más la variable que estimaremos como OTRDOH_B. Por su parte, para construir el total de empleos relativo a estas transferencias (D.7_e) tendremos las *Transferencias periódicas monetarias abonadas a otros hogares* de la ECV, HY130N, junto a la estimación del resto de *Otras transferencias corrientes pagadas* no recogidas en la encuesta, OTPAOH_B.

Finalmente, el saldo de la Cuenta de distribución secundaria de la renta de los hogares, obtenido como diferencia entre los recursos y los empleos de la misma, nos proporciona la *Renta disponible bruta* de este sector (B.6b), que en CNE se calcula como:

Renta disponible bruta (B.6b)

- = Saldo de rentas primarias bruto (B.5b_r)
- Impuestos sobre la renta (D.51_e)
- Imp. sobre el patrimonio y otros impuestos corrientes (D.59_e)
- Cotizaciones sociales netas (D.61_er)
- + Prest. sociales distintas de las prest. sociales en especie recibidas (D.62_r)
- Prest. sociales distintas de las prest. sociales en especie pagadas (D.62_e)
- + Indemnizaciones de seguro no vida (D.72_r)
- + Transferencias corrientes diversas recibidas (D.75_r)
- Primas netas de seguro no vida (D.71_e)
- Transferencias corrientes diversas pagadas (D.75_e)

Y, en la ECV, tal y como hemos analizado a lo largo de este apartado, la Renta disponible bruta de los hogares (B.6b_Bdh) comparable con la de CNE (véase cuadro 2.8), la calcularemos como:

$$\begin{aligned} B.6b_Bdh = & B.5b_Bdh - IRPF - IPATR - CSEFEMP - CSIMEMP - CSEFCP \\ & - CSEFHOG - CSSHOG + OCS + PSDES_B + PSJUB_B + PSSUP_B \\ & + ORNC2_B + PSENF_B + PSINV_B + ORNC3_B + PSEST_B \\ & + ORNC4_B + PSFAM_B + PSVIV_B + PSONC_B - INDNOVIDA1_B \\ & + POSSS_r - POSSS_e + INDNOVIDA1_B + INDNOVIDA2_B \\ & + TPRDOH_B + OTRDOH_B - PRIMNOVIDA_B - TPPAOH_B \\ & - OTPAOH_B \end{aligned}$$

Cuadro 2.8 Cuenta de distribución secundaria de la renta del sector hogares: correspondencia entre la CNE y la ECV

Concepto CNE	Código partida CNE (bruta)	Concepto ECV	Código variable ECV (bruta)
+ Saldo de rentas primarias bruto (recursos)	B.5b_r	= SALDO DE RENTAS PRIMARIAS BRUTO	B.5b_Bdh
- Impuestos corrientes sobre la renta, el patrimonio, etc. (empleos)	D.5_e	= IMPUESTOS CORRIENTES SOBRE LA RENTA, EL PATRIMONIO, ETC.	D.5_e_Bdh
- Impuestos sobre la renta (empleos)	D.51_e	- Impuesto sobre la renta	IRPF=HY140G-CSEFHOG*
- Otros impuestos corrientes (empleos)	D.59_e	- Impuesto sobre el patrimonio y otros impuestos corrientes	IPATR = HY120G
- Cotizaciones sociales netas (empleos-recursos)	D.61_er	= COTIZACIONES SOCIALES NETAS	D.61_er_Bdh
- Cotizaciones sociales efectivas a cargo de los empleadores (empleos)	D.611_e	- Cotizaciones sociales efectivas a cargo de los empleadores	CSEFEMP= $\sum(PY030G)$ - CSEFCP
- Cotizaciones sociales imputadas a cargo de los empleadores (empleos)	D.612_e	- Cotizaciones sociales imputadas a cargo de los empleadores	CSIMEMP*
- Cotizaciones sociales efectivas a cargo de los hogares (empleos)	D.613_e***	- Cotizaciones sociales efectivas a cargo de los hogares	CSEFCP + CSEFHOG*
- Cotizaciones sociales suplementarias a cargo de los hogares (empleos)	D.614_e***	- Cotizaciones sociales suplementarias a cargo de los hogares	CSSHOG*
+ Servicios de los sistemas de seguros sociales (recursos)	D.61SC_r	+ Otras de cotizaciones sociales	OCS*
+ Prestaciones sociales distintas de las transferencias sociales en especie (recursos)	D.62_r	= PRESTACIONES SOCIALES RECIBIDAS	D.62_r_Bdh
		+ Prestaciones por desempleo	PSDES_B= $\sum(PY090G)$
		+ Prestaciones por jubilación	PSJUB_B= $\sum(PY100G)$
		+ Prestaciones por supervivencia	PSSUP_B= $\sum(PY110G)$
		+ Prestaciones de supervivencia percibidas por los menores de 16 años	ORNC2_B*
		+ Prestaciones por enfermedad	PSINF_B= $\sum(PY120G)$
		+ Prestaciones por invalidez	PSINV_B= $\sum(PY130G)$
		+ Prestaciones por enfermedad e invalidez percibidas por los menores de 16 años	ORNC3_B*
+ Prestaciones de seguridad social en efectivo (recursos)	D.621_r***	+ Ayuda para estudios	PSEST_B= $\sum(PY140G)$
+ Prestaciones de asistencia social en efectivo (recursos)	D.623_r***	+ Ayudas para estudios percibidas por los menores de 16 años	ORNC4_B*
		+ Ayuda por familia/hijos	PSFAM_B=HY050G
		+ Ayuda para vivienda	PSVIV_B=HY070G
		+ Ingresos por asistencia social	PSONC_B=HY060G
		- Indemnizaciones de seguros no vida recibidas lesiones o daños sufridos por las personas (incluidas en Prestaciones por enfermedad o invalidez).	INDNOVIDA1_B*
+ Prestaciones de otros sistemas de seguros sociales (recursos)	D.622_r***	+ Prestaciones de otros sistemas de seguros sociales	POSSS_r = CSIMEMP
- Prestaciones sociales distintas de las transferencias sociales en especie (empleos)	D.62_e	= PRESTACIONES SOCIALES PAGADAS	D.62_e_Bdh
- Prestaciones de otros sistemas de seguros sociales	D.622_e	- Prestaciones de otros sistemas de seguros sociales	POSSS_e*
+ Otras transferencias corrientes (recursos)	D.7_r	= OTRAS TRANSFERENCIAS CORRIENTES PERCIBIDAS	D.7_r_Bdh
+ Indemnizaciones de seguro no vida (recursos)	D.72_r	+ Indemnizaciones de seguro no vida que las aseguradoras pagan a los hogares por lesiones o daños sufridos por las personas o los bienes.	INDNOVIDA1_B* + INDNOVIDA2_B*
+ Transferencias corrientes diversas (recursos)	D.75_r	+ Transferencias periódicas monetarias brutas percibidas de otros hogares	TPRDOH_B = HY080G
		+ Otras transferencias corrientes percibidas por los hogares	OTRDOH_B*
- Otras transferencias corrientes (empleos)	D.7_e	= OTRAS TRANSFERENCIAS CORRIENTES PAGADAS	D.7_e_Bdh
- Primas netas de seguro no vida (empleos)	D.71_e	- Primas a pagar en virtud de pólizas de seguro no vida suscritas por los hogares	PRIMNOVIDA_B*
- Transferencias corrientes diversas (empleos)	D.75_e	- Transferencias periódicas monetarias brutas abonadas a otros hogares	TPPAOH_B = HY130G
		- Otras transferencias no periódicas monetarias pagadas por los hogares	OTPAOH_B*
= Renta disponible bruta	B.6b	= RENTA DISPONIBLE BRUTA DE LOS HOGARES	B.6b_Bdh

(*) : Variables no disponibles en la ECV pendientes de estimar en el desarrollo de este trabajo.

(***) : Variables desconocidas en CNE base 2010 debido a las nuevas normas de publicación.

Fuente: Elaboración propia a partir de los datos del INE de la CNE (base 2010) y de las ECV.

4. Niveles de congruencia entre la ECV y CNE

El objetivo de este apartado es calcular y analizar las discrepancias existentes entre la ECV y la CNE, a nivel nacional y para los años de nuestro estudio (2008 y 2011), a partir de las diferencias estudiadas en las cuentas homogéneas anteriormente definidas.

Para ello, en primer lugar, se ha procedido a la creación de una base de datos de hogares (en adelante, BDH) para cada base de la ECV, 2004 y 2013, que toma como principal fuente de información de partida los microdatos de las ECV que contienen los ingresos para los años objeto de estudio 2008 y 2011 (ECV 2009 y ECV 2012, respectivamente). Tras un análisis preliminar, se han llevado a cabo un conjunto de acciones de homogeneización sobre los microdatos de ambos años con el fin de salvar las diferencias debidas a los cambios realizados en la publicación de cada encuesta. Posteriormente, con los microdatos homogeneizados, se han creado las variables necesarias para la construcción de las partidas contables de la CRSH, de acuerdo a la metodología expuesta en los cuadros 2.7 y 2.8.

Con base en los cuadros 2.7 y 2.8, el cuadro 2.9 presenta dichas partidas contables calculadas para los años de estudio, 2008 y 2011, tanto con la BDH que contiene los datos de la ECV en base 2004 como con la construida en base 2013, comparadas con las correspondientes de CNE. Es preciso aclarar que estos cálculos son los realizados en el actual estado de la investigación en este punto; es decir, sin haber estimado aún ninguna de las partidas claramente diferenciadas y específicamente notadas en los párrafos anteriores, que precisan de información complementaria para su estimación y cuya estimación será abordada en los siguientes capítulos, una vez que alleguemos a estas bases de datos información adicional procedente de otras fuentes. En estos cálculos, dichas variables están puestas a cero.

Como se puede observar en dicho cuadro (2.9), las subestimaciones que producen las ECV a nivel agregado para la renta disponible son el fruto de otras subestimaciones que subyacen también a nivel de partidas. Así, los grados de subdeclaración u ocultación de ingresos son bastante diferentes de unas partidas a otras de la ECV, con respecto de la CNE, y dependiendo de la base de la encuesta con la que se realice la comparación. En concreto, en la BDH base 2004, sus errores relativos con respecto a CNE oscilan entre un significativo 95,8% para las *Otras transferencias corrientes recibidas*, en 2011, y un valor mínimo del 16,3% para el *Excedente de explotación neto/ Renta mixta neta* en el mismo año. Por su parte, en la BDH base 2013 vemos cómo la incorporación de los registros administrativos en la producción de los datos de ingresos relativos al hogar en las ECV base 2013 contribuye considerablemente a la reducción de dichos errores con respecto a la CNE en todas las partidas que forma parte de la Renta disponible de los hogares, a excepción de las *Otras transferencias corrientes*, tanto *recibidas* como *pagadas*, situándose también los errores mínimos en la partida *Excedente de explotación neto/ Renta mixta neta*, en 0,5% para el año 2008 y en 2,4% en 2011.

Las diferencias tan amplias de coberturas entre partidas, manifiestan paralelamente distintos grados de ocultación de rentas en las mismas; y este

hecho, podría estar sesgando a la baja las estimaciones de las rentas de aquellos colectivos que se caracterizan por alimentarse principalmente de los tipos de renta en los que más se oculta. Esta asimetría observada en la distribución del desajuste es una prueba más de la necesidad de estimar una distribución personal de la renta corregida que tenga en cuenta dicha asimetría.

Cuadro 2.9
CUENTA DE RENTA DE LOS HOGARES: ECV vs. CNE

			Excedente explotación neto/ Renta mixta neta	Remuneración asalariados	Rentas propiedad (rec.-emp.)	Impuestos corr. sobre la renta, patrimonio, etc. + Cotizaciones sociales netas (emp.-rec.)	Prestaciones sociales distintas transferencias sociales especie (rec.-emp.)	Otras transferencias corrientes (recursos)	Otras transferencias corrientes (empleos)	Renta disponible neta
2008	BDH base 2004	miles de €	98.157.521	434.759.297	12.430.665	153.594.728	113.851.048	2.273.760	4.200.439	503.677.123
	BDH base 2013		127.628.607	488.742.659	31.180.885	191.195.022	126.662.474	2.263.734	4.355.668	580.927.669
	CNE		128.323.000	560.517.000	36.796.000	248.043.000	154.203.000	51.565.000	54.594.000	628.767.000
	Errores relativos BDH b2004/CNE (%)		-23,5%	-22,4%	-66,2%	-38,1%	-26,2%	-95,6%	-92,3%	-19,9%
	Errores relativos BDH b2013/CNE (%)		-0,5%	-12,8%	-15,3%	-22,9%	-17,9%	-95,6%	-92,0%	-7,6%
2011	BDH base 2004	miles de €	107.320.728	386.050.518	9.746.443	140.526.279	136.341.175	1.984.742	3.486.852	497.430.474
	BDH base 2013		125.190.519	451.812.669	26.030.982	185.796.945	150.202.032	1.978.490	3.571.811	565.845.936
	CNE		128.217.000	531.876.000	38.756.000	236.048.000	183.674.000	47.374.000	54.811.000	639.038.000
	Errores relativos BDH b2004/CNE (%)		-16,3%	-27,4%	-74,9%	-40,5%	-25,8%	-95,8%	-93,6%	-22,2%
	Errores relativos BDH b2013/CNE (%)		-2,4%	-15,1%	-32,8%	-21,3%	-18,2%	-95,8%	-93,5%	-11,5%

Fuente: Elaboración propia a partir de los datos del INE de la CNE Base 2010 y los microdatos de las ECV (base 2004 y base 2013).

Asimismo, sabemos que parte de las discrepancias observadas entre ambas fuentes son imputables a las diferencias existentes en las definiciones que hemos analizado en el apartado 2.3 de este capítulo. Por su parte, la distinta población a la que hace referencia cada fuente (45.983.168 y 46.736.256 de personas en 2008 y 2011, respectivamente, en CNE, frente a los 45.870.296 y 46.354.779 de personas de la ECV) explica una parte de dicho sesgo a la baja que se manifiesta en la encuesta a nivel de cada componente de la renta y para la totalidad de ésta.

Por lo tanto, para construir una distribución de la renta a nivel de microdatos más cercana a la realidad que dibuja la CNE que la que se extrae directamente de la información recogida en la ECV, es preciso seguir trabajando en dos líneas: por un lado, en la estimación de las partidas no disponibles en la ECV; y, por otro, en la estimación de la infradeclaración de los hogares en las distintas partidas. Ambas líneas aconsejan utilizar otra información de los hogares que no existe en la ECV, pero de la que se dispone, al menos parcialmente, en otras fuentes (Encuesta de Presupuestos Familiares y Encuesta Financiera de las Familias) y que estudiaremos en el siguiente capítulo. Todo ello con el objetivo final de analizar la distribución personal de la renta que se derive al tener en cuenta estas estimaciones.

Capítulo III

El consumo y la situación financiera de los hogares de las ECV

1. Introducción

En el capítulo anterior hemos llevado a cabo el estudio de dos de las fuentes más importantes que proporcionan información sobre los ingresos de los hogares, tanto a nivel de microdato (ECV) como a nivel agregado (CNE), lo que nos ha permitido poner de manifiesto la infraestimación de la Renta Disponible de los hogares en la ECV para el total nacional y para cada una de las partidas que constituyen la Cuenta de Rentas del Sector Hogares (CRSH), tomando como marco de referencia la CNE. Una vez detectadas estas incongruencias a nivel agregado nos planteamos, a continuación, cómo podríamos repartirlas a nivel de hogar para obtener una distribución personal de la renta corregida que sea consistente con la Renta Disponible del sector hogares que se desprende de la CNE y con cada una de sus partidas.

Habida cuenta de que el ingreso familiar es un flujo que permite consumo y contribuye a los cambios en la riqueza del hogar o del patrimonio neto, el objetivo que nos planteamos en este capítulo consiste en indagar en la literatura y la teoría económica sobre este tipo de relaciones a nivel de microdato. De esta forma, estaremos en condiciones de exponer las relaciones teóricas que deberían darse entre ingresos, gastos y patrimonio neto de las personas y hogares, así como las repetidas contradicciones al respecto empíricamente observadas en las fuentes que los recogen, para concluir con la conveniencia y posibilidad de usar los datos de consumo y de riqueza para corregir parte de la sub-declaración de ingresos que se produce en personas y hogares.

Así pues, dados los ingresos de cada hogar por medio de la ECV, podemos completar este círculo con los estudios sobre consumo -Encuesta de Presupuestos Familiares (EPF)- y riqueza -Encuesta Financiera de las Familias (EFF)- existentes a nivel de hogar con la finalidad de utilizar dichas relaciones teóricas para explicar la ocultación detectada a nivel nacional para cada uno de los hogares. Para ello, procederemos a compilar toda esta información para cada hogar por medio de técnicas de fusión estadística con el fin de obtener un registro de hogares que contenga información económica lo más plausible posible que nos permita poner de manifiesto cómo ocultan los hogares en función de toda esta información complementaria que vamos a fusionar.

Cabe destacar que el marco integrado de microdatos sobre ingresos, consumo y riqueza que presentaremos en este capítulo tiene muchos puntos en común con el marco para los datos macro desarrollados en el Sistema europeo de cuentas económicas integradas (SEC) que subyace en la CNE. Sin embargo, mientras el SEC 2010 se caracteriza por mantener la coherencia en la mayoría de los aspectos que se mantiene para el conjunto de la economía, nuestro estudio a nivel microeconómico tiene una perspectiva referida únicamente a los hogares, no a toda la economía. También difiere del SEC porque contiene un enfoque en el que se reflejan las variaciones en el bienestar económico dentro del sector de los hogares, mientras que el SEC se centra en suministrar las mediciones de las macromagnitudes de ingreso, consumo y riqueza para toda la economía en su conjunto.

2. Estudio de relaciones teóricas para el análisis de la ocultación

Tal y como anunciábamos en la parte introductoria de este capítulo, sabiendo que el ingreso familiar es un flujo que permite consumo y contribuye a los cambios en la riqueza o del patrimonio neto del hogar, para desglosar a nivel de microdato la ocultación de ingresos del sector hogares detectada en la ECV -fruto de su comparación con la CNE a nivel nacional y para cada una de las partidas de la Cuenta de Rentas de dicho sector- se hace primordial poder hacer uso de información complementaria sobre el patrón de consumo y riqueza que siguen dichos hogares. Para ello deberemos, en primer lugar, estudiar en profundidad las relaciones que se esconden entre estas tres magnitudes según la Teoría Económica y la literatura existente.

Por tanto, este apartado lo dedicaremos a investigar sobre los conceptos de consumo y riqueza y los nexos que existen entre éstos y la renta en el ámbito del sector hogares.

Cristóbal (2010) define el consumo como la parte del gasto en bienes y servicios que no está destinada a una actividad productiva o a la generación de otros bienes y servicios. Por esta vía, estaríamos definiendo la variable en función de la aplicación que se le da a los bienes y servicios, los cuales pueden clasificarse en bienes y servicios de inversión y bienes y servicios de consumo. La diferencia entre ellos radica en que, mientras los primeros se utilizan en la producción, los segundos se agotan una vez consumidos. Esta clasificación no siempre es estricta y da lugar a la aparición de zonas más o menos grises. Lo que es un bien de inversión para determinada persona, puede ser un bien de consumo para otra, en función de la finalidad que persiga con su aprovechamiento. Por ejemplo, un automóvil puede ser un bien de consumo para una familia, y un bien de inversión si lo emplea un comercio para reparto de sus productos a domicilio.

En concreto, la fuente de información que tomamos como referencia más fiable para la medición de las magnitudes económicas, esto es, la CNE, distingue, en función de las unidades que realicen el gasto, el gasto en consumo final de los hogares, el de las instituciones sin fines de lucro al servicio de los hogares (ISFLSH) y el de las administraciones públicas (AA.PP.).

Según la CNE, el gasto en consumo final de los hogares incluye, fundamentalmente, las compras de bienes y servicios realizadas por aquellos para satisfacer sus necesidades habituales. Estas compras pueden realizarse tanto en el territorio económico como en el resto del mundo. Las cuentas nacionales desagregan este gasto en función de su finalidad, es decir, del tipo de bien o servicio adquirido, utilizando para ello la clasificación COICOP³⁶. Así se distinguen doce grandes categorías de gasto (alimentos, artículos de vestir, vivienda, salud, transporte, etc.). Además de dichas compras, el gasto en consumo final de los hogares incluye también una serie de gastos imputados, que las personas pueden no ser conscientes de haber realizado, pero que se

³⁶ La clasificación COICOP (abreviatura del inglés Classification Of Individual Consumption by Purpose) es una clasificación del gasto de consumo realizado por los hogares desarrollada por la División de Estadística de Naciones Unidas. En el menor nivel de desagregación (clasificación de un dígito), existen doce tipos de bienes y servicios.

consideran como tales para dar una mayor consistencia a las cuentas económicas. Entre ellos, destacan los siguientes, de los que ya hablamos en el Capítulo 1:

- Los servicios de alquiler de las viviendas ocupadas por sus propietarios. La adquisición de una vivienda (bien de inversión), permite al hogar propietario realizar una producción de servicios de alquiler, que es consumida directamente por el propio hogar.
- La renta en especie. Los asalariados pueden percibir bienes y servicios de forma gratuita o a un precio muy reducido como parte de su salario (comidas, uso de vehículos o de viviendas, etc.). Esos ingresos en especie formarían parte de su remuneración final y, por lo tanto, aumentarían en la misma cuantía el gasto en consumo final en las categorías de gasto correspondientes.
- Los servicios de intermediación financiera medidos indirectamente (SIFMI). Los depósitos y préstamos que los hogares tienen con las entidades financieras generan un gasto en servicios de intermediación que se estiman indirectamente a partir de los intereses recibidos o pagados a los hogares.

Desde una perspectiva geográfica, el gasto en consumo final de los hogares puede medirse también para todas las unidades que lo realizan en el territorio económico, tanto residentes como no residentes (por ejemplo, en este último caso, el gasto en consumo de los turistas). De esta forma, las cuentas económicas hacen referencia también al gasto en consumo final interior (cuadro 3.1).

Cuadro 3.1
**EL GASTO EN CONSUMO FINAL DE LOS HOGARES DESDE
UNA PERSPECTIVA GEOGRÁFICA**

Concepto	Componentes
Gasto en consumo final de los hogares	Gasto en consumo de los hogares residentes realizado en España
	Gasto en consumo de los hogares residentes realizado en el resto del mundo
Gasto en consumo final de los hogares interior	Gasto en consumo final de los hogares residentes realizado en España
	Gasto en consumo final de los hogares no residentes realizado en España

Fuente: Instituto Nacional de Estadística.

Desde una óptica analítica, los expertos prefieren otra clasificación de los gastos en consumo final de los hogares. En concreto, se inclinan a que una división en bienes y servicios, junto a una desagregación de los bienes con respecto a su durabilidad (duraderos, semiduraderos y no duraderos), es más útil para el análisis del comportamiento de los hogares frente al consumo, dado que aquellos tienen una pauta de decisión diferente en cada uno de esos tipos de gasto.

Por su parte, las ISFLSH son unidades de la economía, que se consideran productores privados fuera del mercado, son financiadas por los hogares y

otras unidades, y suministran su producción a aquéllos en forma de servicios de distintos tipos (partidos políticos, sindicatos, asociaciones de consumidores, asociaciones religiosas, algunas organizaciones no gubernamentales, etc.). Estas unidades también realizan gasto en consumo final, siendo los hogares los únicos beneficiarios de dicho gasto.

De la misma forma, las AA.PP. también realizan gasto en consumo final, gasto del que se benefician tanto los hogares como el resto de la sociedad. En CNE este gasto se desagrega en función de su finalidad utilizando la COFOG (Classification of the Functions of Government) que diferencia entre diez funciones de gasto para las AA.PP.

Finalmente, la Contabilidad Nacional define también el concepto de consumo final efectivo, ligado no con el sector que realiza el gasto en consumo sino con el que se beneficia finalmente de dicho gasto. El consumo final efectivo puede ser individual o colectivo, en función de si se puede identificar claramente al beneficiario de dicho consumo (individual) o bien se destina a toda una comunidad de forma simultánea (colectivo).

Cuadro 3.2
EL GASTO EN CONSUMO FINAL Y EL CONSUMO FINAL EFECTIVO

Sector que realiza el gasto				
Tipo de consumo	Hogares	ISFLSH	AA.PP.	Total
Consumo individual	Gasto en consumo final de los hogares	Gasto en consumo final de las ISFLSH (Transferencias Sociales en Especie a hogares)	Gasto en consumo final individual de las AA.PP. (Transferencias Sociales en Especie a hogares)	Consumo final efectivo individual
Consumo colectivo	-	-	Gasto en consumo final colectivo de las AA.PP. (Transferencias Sociales en Especie a hogares)	Consumo final efectivo colectivo
Total	Gasto en consumo final de los hogares	Gasto en consumo final de las ISFLSH	Gasto en consumo final de las AA.PP.	Consumo Final Efectivo (Total economía) = Gasto Consumo Final (Total economía)

Fuente: Elaboración propia a partir del SEC 2010.

De esta manera, de forma individual, los hogares habrán podido consumir, por un lado, todos los bienes y servicios que han adquirido directamente (su gasto en consumo final) y aquella parte del gasto de las ISFLSH y de las AA.PP. que se suministra a los hogares sin ningún tipo de pagos a través de lo que se denominan, en Contabilidad Nacional, transferencias sociales en especie. En el caso de las ISFLSH se transfiere a los hogares individualmente todo su gasto en consumo y en el caso de las AA.PP. aquella parte que corresponde al gasto en el que se puede identificar al beneficiario (educación y sanidad públicas, asistencia social, etc.). El resto del gasto de las AA.PP. correspondiente a los servicios que se prestan de forma colectiva (policía, defensa, etc.) será el consumo final colectivo efectivo (de las AA.PP.). El total de gasto en consumo y el total de consumo efectivo deben ser iguales. En el cuadro 3.2 se presentan

las relaciones descritas entre el gasto consumo final y consumo final efectivo por sectores y para el total de la economía.

En relación al sector de estudio que nos ocupa a lo largo de este trabajo, nuestro interés se va a centrar en el consumo privado de las familias o economías domésticas. El consumo privado es el componente más importante de la demanda agregada de un país y, por ello, resulta primordial a la hora de estudiar la evolución de una economía, tal y como expresa Jeffrey Sachs (citado en Larrain & Sachs, 2002, p.77): “el efecto acumulativo de las decisiones en el consumo y ahorro de las familias contribuye a determinar la tasa de crecimiento de la economía, la balanza comercial y el nivel de producto y empleo”.

Lo anterior se fundamenta en que el Producto Interior Bruto o Renta Nacional de un país consiste, básicamente, en la suma del consumo privado, la inversión privada, el gasto público y el saldo exterior neto. Como el consumo privado es la parte proporcional más importante de dicho sumando, las variaciones del consumo tienen consecuencias a corto plazo muy importantes en las variaciones de la producción, la renta, el empleo y, por tanto, son determinantes fundamentales de los ciclos económicos (UN, 2008).

Por otro lado, toda la renta disponible que no se consume se ahorra. El ahorro es la base de la inversión y de la acumulación del capital y, por tanto, del crecimiento a largo plazo. Así, los factores que afectan al ahorro son los que determinan el crecimiento futuro y la prosperidad de un país. Los economistas muestran un enorme interés sobre las variaciones del consumo a corto plazo y sobre las variaciones del ahorro (o el no-consumo), a largo plazo, ya que ambos son factores determinantes, en distintos espacios temporales, del desarrollo de la economía (Mankiw, 2010). Debido al periodo de estudio investigado en este trabajo, nos centraremos fundamentalmente en las repercusiones del consumo de las familias en el corto plazo.

En la economía española, el consumo privado explica una proporción sustancial de las fluctuaciones de la actividad económica en el ciclo más reciente, dado su elevado peso dentro del PIB -en torno al 57% en 2015, en términos nominales- y la propia intensidad de las oscilaciones que este componente de la demanda ha experimentado durante ese período.

Una aproximación al concepto de consumo privado que se desprende de forma agregada de la CNE nos la proporcionan a nivel de microdato los datos de gasto en consumo de los hogares residentes en España que se publican en la Encuesta de Presupuestos Familiares (EPF) llevada a cabo por el INE.

La selección de la EPF como fuente de referencia para conocer los gastos de consumo de los hogares se explica por ser la que ofrece microdatos más detallados de consumo a nivel familiar, a la vez que de renta y una rica información sobre características geográficas, demográficas y socioeconómicas de los hogares entrevistados, junto con otras informaciones igualmente valiosas en torno al endeudamiento, las características de la vivienda y el equipamiento del hogar o el consumo no monetario. Por otra parte, multitud de estudios avalan la tesis de que cuando se recogen simultáneamente ingresos y gastos para los hogares, los datos de gastos suelen

ser más fiables y mayores que los ingresos declarados³⁷. En este caso, la posibilidad de comparar dichas variables nos permitirá analizar la congruencia de los datos recogidos en cada hogar y proceder a la corrección de los ingresos, generalmente al alza, con la finalidad de mejorar la consistencia global de la información.

2.1. Determinantes del consumo y el ahorro

Los economistas se han interesado desde siempre por los factores que determinan la proporción en que la sociedad distribuye su renta entre consumo y ahorro. Tradicionalmente la investigación teórica y empírica de estos elementos se ha basado en el concepto de función de consumo, que recoge una serie de variables que influyen sobre el consumo, así como el carácter y magnitud de sus efectos. La renta es, evidentemente, una de las más destacadas variables a considerar, por lo que gran parte de las investigaciones se han ocupado del estudio de la dependencia entre ésta y el consumo.

En este contexto, aparece el ahorro y lo hace como diferencia, positiva o negativa, entre renta y consumo. Así, todas las relaciones incorporadas a la función de consumo pueden asimismo incorporarse a la función de ahorro. Resulta indiferente que la atención se centre sobre los determinantes del consumo o sobre los del ahorro ya que las conclusiones de cualquiera de las orientaciones pueden aplicarse al otro. Por ejemplo, la afirmación de que las economías domésticas españolas gastan globalmente en consumo el 91% de su renta después de los impuestos equivale a decir que ahorran el 9% de su renta disponible. En todo caso, cualquiera de estos dos planteamientos no es más que una primera aproximación a la realidad.

La fuente que utilizaremos para obtener información sobre el ahorro de los hogares españoles y cada uno de sus miembros es la Encuesta Financiera de las Familias (EFF), que como ya explicamos en el Capítulo 1, elabora el Banco de España desde el año 2002 con una frecuencia trienal y contiene una gran cantidad de información sobre la situación patrimonial y financiera de las familias, permitiendo de este modo relacionar para un mismo hogar las características demográficas, las rentas, los activos reales y financieros, las deudas y el gasto.

Para entender el comportamiento del consumo y sus fluctuaciones (y, por ende, del ahorro), es necesario saber cuáles son sus factores determinantes y cómo forman las unidades de consumo sus decisiones en cuanto al mismo. Para ello, a continuación, se enumerarán una batería de factores que influyen en el consumo partiendo de la teoría Keynesiana, y teniendo en cuenta también, entre otros, los aportes teóricos de Duesenberry, Modigliani y Friedman.

³⁷ Sanz (1995) destaca la infravaloración de las rentas en la Encuesta de Presupuestos Familiares de 1990-1991 respecto de la CNE de 1990. Por su parte, Ruiz-Castillo (1987) resalta lo llamativo de que un 60% de los hogares declare un ingreso menor que su gasto en la Encuesta de Presupuestos Familiares de 1980-1981, mientras que Ruiz-Huerta y Martínez (1994) y Cantó (1997) detectan sesgos similares en la Encuesta de Presupuestos Familiares de 1990-1991 y en la Encuesta Continua de Presupuestos Familiares, respectivamente.

Siguiendo a Ackley (1965)³⁸, se pueden encontrar diversos tipos de factores que se asocian al consumo dentro de la teoría económica: objetivos, subjetivos, estructurales y políticos³⁹.

Los *factores objetivos* hacen referencia a variables económicas capaces de experimentar una apreciable variación a corto plazo y que pueden ser medidas a escala cuantitativa, de forma que se podría permitir su incorporación a una función de consumo estadística. Sería de esperar que algunos de estos factores objetivos influyeran en el comportamiento individual respecto del consumo en general; otros, estarán sobre todo relacionados con el gasto en bienes duraderos.

Por su parte, los *factores subjetivos o psicológicos*, incluyen elementos de expectativas y actitudes de los individuos frente al consumo.

En tercer lugar, como *factores estructurales*, estudiaremos algunos aspectos pertinentes al problema de la agregación. Aunque las propensiones al consumo de grupos o tipos específicos de hogares podrían ser estables, un cambio en las ponderaciones relativas de los diversos grupos puede alterar la relación analizada entre el consumo y el ingreso, u otros factores explicativos.

Finalmente, los *factores de carácter político* afectan claramente al papel del consumo en la economía por medio, sobre todo, de su influencia en el tamaño del ingreso disponible (variable de ingreso que utilizaremos para hacer referencia al comportamiento individual) en relación con la producción total. El más importante de éstos es la política fiscal.

A continuación, vamos a analizar de forma pormenorizada estos factores tratando de apoyarnos en estudios existentes en la literatura y teoría económica que justifiquen su relación con el consumo o el ahorro. Al mismo tiempo, trataremos de identificar en qué fuente de microdatos, de entre las que investigamos, podemos localizar dicha información, y a través de qué variables. La localización de estas variables será clave para llevar a cabo, mediante ellas, la fusión estadística de los datos de la ECV -presentes en nuestra BDH- con los de la EPF en primer lugar y, posteriormente, con los de la EFF, lo que nos permitirá disponer de una gran base de datos que contendrá información sobre ingresos, gastos y variables de riqueza, a través de la cual podremos proceder a la comprobación de las relaciones teóricas estudiadas y detectar la posible ocultación de ingresos a nivel de hogar.

Debido a la ausencia de información de carácter subjetivo en las fuentes de microdatos utilizadas en este trabajo y al periodo de estudio que se analiza en el mismo, no tendremos en cuenta, por motivos obvios, ni los factores subjetivos o psicológicos ni los políticos.

³⁸ *Teoría Macroeconómica*, G. Ackley, UTEHA, 1965.

³⁹ Keynes, que nunca atribuyó importancia causal exclusiva al ingreso, estudió otros factores bajo los encabezamientos "objetivos" y "subjetivos". Ackley emplea una clasificación algo más detallada, basada solo en parte en la distinción de Keynes.

2.1.1. Factores objetivos

INGRESO REAL

J. M. Keynes en su *General Theory of Employment, Interest and Money*, de 1936, fue el primer economista en describir de forma plena y clara que, en toda economía en la que la gente tenga libertad de elección, el volumen total de gastos de consumo personal en un año dado está determinado principalmente por el volumen de renta disponible recibido por la gente en ese año. Esta proposición, sencilla de significado y aparentemente correcta, es la piedra angular de la *Teoría de la demanda de consumo*. Sin embargo, es fácil hallar con anterioridad otros autores que habían estado cerca de enunciar la misma idea. A pesar de ello, la "función de consumo" es considerada apropiadamente como una invención keynesiana porque descansa en el corazón del sistema teórico de Keynes.

El enfoque keynesiano no tuvo en cuenta aspectos tales como la riqueza, los gustos y los precios relativos como factores determinantes. Le importaba únicamente el corto plazo, con lo cual, a pesar de haber observado otros posibles determinantes, no los incluyó en su análisis.

Para estudiar la forma en que los gastos de consumo están relacionados con la renta disponible, podemos empezar por suponer que tales gastos varían directamente con la renta disponible. En segundo lugar, podemos concretar más y decir algo sobre la cantidad en que varían tales gastos al variar la renta. Así lo hizo Keynes (1936) al establecer su *ley psicológica fundamental* de que, "cuando aumenta la ocupación aumenta también el ingreso global real de la comunidad; la psicología de ésta es tal que cuando el ingreso real aumenta, el consumo total crece, pero no tanto como el ingreso". En otras palabras, a medida que la renta aumenta, los consumidores gastarán parte de ese aumento, pero no todo, pues preferirán ahorrar parte. Así, el aumento total de renta será igual a la suma del aumento de los gastos de consumo y del aumento del ahorro personal, si suponemos para mayor sencillez, que el interés pagado por los consumidores es cero. A esta ley psicológica la llamó la *propensión marginal a consumir*, siendo una característica de ésta su estabilidad.

Finalmente, ¿podemos ser más concretos respecto a la naturaleza de esta relación? Keynes tenía gran confianza en que su *ley psicológica fundamental* era correcta, pero añadió, con menos confianza, el argumento de que a medida que la renta aumenta se consumirá una menor proporción de renta (o se ahorrará una mayor proporción). Si esto es cierto, ello significa que al aumentar la renta no sólo aumentará la relación entre ahorro y la renta. Keynes creía que era de esperar que esto ocurriría como norma general, pues a pesar de que "la satisfacción de las necesidades primarias inmediatas de un hombre y de su familia es normalmente un motivo más fuerte que el que lleva a la acumulación", este último "cambia de forma importante cuando ya se posee un cierto margen de comodidades".

Esta relación entre consumo e ingreso disponible, a través de un concepto típicamente keynesiano, como la *propensión marginal a consumir*, juega un

papel fundamental no solo en la determinación del consumo sino en el multiplicador y, por lo tanto, en la determinación del nivel de la renta.

Podemos representar la función consumo keynesiana de la siguiente manera:

$$C = C_0 + cY_d$$

donde C_0 es una constante positiva que representa el consumo autónomo, es decir, aquél que se produce independientemente del nivel de ingreso Y_d , y c constituye la propensión marginal a consumir que es una constante positiva y menor que la unidad.

En definitiva, los postulados principales de esta teoría son que los consumidores destinan una parte de todo aumento de renta al ahorro y ahorran una parte mayor de una renta grande que de una pequeña. Esta teoría específica del gasto de consumo, aparte de algunas simplificaciones de carácter no básico, es, esencialmente, la tesis que Keynes adelantó en su *Teoría General*.

La teoría de Keynes ha sido elaborada por otros y ha llegado a ser conocida con el nombre de la *teoría o hipótesis de la renta absoluta*. Se han desarrollado también teorías contrarias, siendo las dos más importantes las *teorías o hipótesis de la renta permanente* y el de la *renta relativa*. Cada una de estas tres teorías del gasto de consumo implica una teoría concreta, pero distinta, de conducta del consumidor, es decir, cada una asigna a la renta un papel diferente en cuanto a su influencia sobre los gastos de consumo, o, dicho de otra forma, cada una ve a la renta, como determinante del consumo, de forma diferente. Las estudiamos a continuación.

INGRESO REAL RETARDADO

Como se ha señalado, Keynes concibió la función de consumo como dependiente de la renta disponible, es decir, de la renta nacional excluidos los impuestos directos, de tal foma que, a mayor renta disponible mayor consumo. Más adelante, Duesenberry profundizó el análisis haciendo depender la demanda de consumo tanto de la renta anterior como de la renta corriente, con su hipótesis sobre la "renta relativa". Para Duesenberry el consumo corriente no solo depende de la renta corriente sino también de la historia de la renta anterior. Los individuos basan su consumo en los niveles de renta pasados más altos y ello lleva consigo a que exista un retraso temporal entre demanda de consumo y renta, ya que el consumo se adapta con atraso a las variaciones de renta, ya que una vez que la renta se reduce el consumo muestra una inercia a caer.

Así pues, el supuesto que sostiene el modelo anterior en el que el consumo del período t depende de la renta del mismo período, ha sido discutido basándose en la existencia de motivos psicológicos, técnicos e institucionales por los que las variaciones en la renta no provocan variaciones en el consumo sino después de un cierto retardo temporal. Por lo tanto, cabe postular que el consumo depende no sólo de la renta percibida en ese año, sino que también influye la renta percibida en periodos anteriores. En Barbancho (1961) se expone dicha relación:

$$C_t = \beta_0 + \beta_1 Y_t + \beta_2 Y_{t-1} + u_t$$

siendo $\beta_0 > 0$, $0 < \beta_1 < 1$ y $\beta_2 > 0$.

Cabe destacar que, en lo referente al consumo, los retardos superiores a un año carecen prácticamente de significación. En este sentido es ilustrativo el ensayo de Fisher, quien al relacionar el consumo con la renta retardada obtiene los siguientes coeficientes de correlación: $r=0,88$ para un retardo de tres meses; $r=0,73$ cuando el retardo es seis meses, $r=0,41$ para el caso de nueve meses, y $r=0,20$ para un año de retardo.

TIPO DE INTERÉS

Podría pensarse a priori que los consumidores racionales ahorrarán más a tipos de interés altos que a bajos. Un interés alto estimulará postergaciones en el consumo. Asimismo, la disminución en el interés parece estimular ciertas compras financiadas.

Tradicionalmente, la teoría económica ha venido señalando la existencia de efectos contrapuestos de los tipos de interés sobre el ahorro de las familias, de manera que el signo que cabe esperar de esta relación es indeterminado. En concreto, hay un efecto renta, de acuerdo con el cual una elevación de los tipos de interés aumenta el valor esperado de las rentas derivadas de la riqueza, lo que permite mantener el consumo futuro con un volumen de ahorro menor. Por tanto, el efecto renta es negativo, de manera que alzas en el tipo de interés reducen el ahorro. No obstante, se produce, asimismo, un efecto de sustitución entre consumo presente y futuro: dado que esta elevación disminuye el coste presente del consumo futuro, el efecto sustitución desalienta el consumo presente del individuo, estimula el futuro, aumentando así el ahorro. El efecto total de ambos resulta a priori ambiguo y depende de si domina uno u otro. Por eso se ha sostenido, a menudo, que la cuestión es eminentemente empírica (Gylfason, 1981; Smith, 1989).

Trabajos teóricos más recientes, como el de Zabalza y Andrés (1991) o Argimón *et al.* (1992), muestran que en el corto plazo los tipos reales de interés tienen una incidencia pequeña pero significativa sobre el ahorro familiar de signo positivo, de tal manera que aumentos en los tipos generan aumentos en el ahorro (y, por ende, reducciones en el consumo).

En definitiva, a pesar de tratarse de una idea sugerida por el propio Keynes, y aún habiéndose constatado dicha incidencia positiva pequeña, pero significativa, del tipo de interés sobre el ahorro familiar a nivel agregado, la incidencia concreta que tiene sobre cada uno de los hogares a nivel individual es mucho más variable y compleja, dependiendo de otros muchos factores de difícil control, razón por la cual, junto al carácter transversal y microeconómico de nuestro estudio, no la tendremos en cuenta a lo largo de este trabajo.

FACILIDADES PARA ACCEDER A LA FINANCIACIÓN

Las fáciles condiciones de crédito ayudan a que la necesidad, o prisa, por comprar, se convierta pronto en una realidad.

En Murcia (2007), el acceso al crédito se enfoca como un servicio financiero importante que demandan los hogares. Un mayor acceso a éste le permite a un conjunto más amplio de personas suavizar su consumo de manera

intertemporal incrementando así su bienestar y unas mejores condiciones de crecimiento económico. En este estudio se encontró que la probabilidad de acceder a un crédito depende en gran medida de las características socioeconómicas de los hogares. Una mayor capacidad de generar ingresos en el futuro la cual está representada por unos mayores niveles de ingreso, riqueza, la presencia de un contrato de trabajo y unos mayores años de educación incrementan de manera significativa las probabilidades de contar con los servicios analizados. Adicionalmente, la posición geográfica afecta en gran medida a la probabilidad de acceder al crédito (menor en las zonas rurales).

En concreto, en ese mismo trabajo, para el caso del crédito hipotecario, pertenecer al quintil más alto de riqueza aumenta la probabilidad de ser usuario de crédito en un 11,7% en comparación a un hogar que se encuentra en el quintil de riqueza más bajo. Para los créditos con tarjeta de crédito, la concentración en los niveles de riqueza altos es mucho mayor, puesto que pertenecer al quintil más alto de riqueza incrementa en un 27,1% la probabilidad de ser usuario de crédito, en comparación a un hogar en el menor quintil de riqueza.

GANANCIAS DE CAPITAL

Keynes dio reconocimiento especial a la posibilidad de que el gasto de consumo pudiera ser influido no sólo por el ingreso, sino por las ganancias de capital. En verdad, especuló con que el auge repentino del consumo a finales de los años '20 en Estados Unidos podría haber sido originado en las ganancias de capital que estaban haciéndose en el mercado de valores. Pigou ha tratado de incluir las ganancias de capital en una función de consumo estadística. Esto conlleva un reconocimiento indirecto de que el consumo puede ser influido por la *riqueza*. En Guiso *et al.* (2005) se deduce que la propensión marginal a consumir a partir de una ganancia de capital es 0,019, sólo ligeramente menor que la propensión marginal a consumir de la riqueza total de inicio de período (0,02).

RIQUEZA

El origen de la investigación del efecto *riqueza* se encuentra en el conjunto de las innovaciones teóricas de los años treinta. Keynes (1936) fue el primer teórico de relevancia que postuló explícitamente la existencia de una relación efectiva entre la riqueza financiera y las decisiones de consumo. Las aportaciones posteriores por parte de Haberler (1939), Pigou (1943) y Patinkin (1956) acabaron por concretar esa relación: los saldos monetarios reales de las familias pueden provocar variaciones en el nivel del consumo agregado. Este fue el referente teórico básico hasta que, años más tarde, Modigliani y sus colaboradores extendieron el alcance del análisis del llamado efecto *Pigou* estableciendo el punto de partida de todas las investigaciones posteriores. El análisis del efecto que los cambios de la riqueza, y no sólo de su componente financiera, proyectan sobre el consumo, constituye el marco general de los trabajos ulteriores relacionados con el efecto *riqueza*.

También se engloban dentro de la investigación teórica relativa al efecto *riqueza* las aportaciones de Brumberg y Modigliani (1954) y de Friedman (1957), conocidas como modelos de renta a largo plazo.

La *hipótesis del ciclo vital* desarrollada por Modigliani y Brumberg (1954) sostiene que los individuos planifican su consumo y su ahorro, a lo largo de toda su vida, con el fin de obtener una senda de consumo estable y creciente. Consecuentemente, esta teoría considera que el consumo de un individuo depende de la renta que percibe a lo largo de toda su vida, eso sí, subrayando que ésta varía sistemáticamente a lo largo del tiempo, y que el ahorro, tanto positivo como negativo, es utilizado como instrumento a través del cuál los individuos obtienen el nivel de consumo planeado.

Por su parte, la *hipótesis de la renta permanente* formulada por Friedman (1957) sugiere que la renta de los individuos experimenta variaciones aleatorias y temporales de un año a otro. De este modo, el flujo de ingresos percibido por un individuo puede ser desglosado entre renta permanente y renta transitoria. La renta permanente puede ser definida como aquella proporción de la renta que los individuos esperan mantener en el horizonte temporal considerado, y engloba a la riqueza humana, entendiendo como tal el capital humano, y a la riqueza no humana, tanto inmobiliaria como financiera. La renta transitoria refleja las fluctuaciones inesperadas de la renta actual. De manera análoga, el consumo realizado por el individuo puede dividirse entre consumo permanente y consumo transitorio. El primero es definido como el valor del consumo esperado en la compra de bienes corrientes y de servicios; el segundo expresa el consumo no previsto. Consecuentemente, esta teoría sugiere una función de consumo determinada exclusivamente por la renta permanente, destinándose la renta transitoria al ahorro.

La disponibilidad de datos adecuados sobre la riqueza agregada de las familias americanas posibilitó, en la década de los sesenta, la investigación empírica (Gil, 2003). En 1963, Ando y Modigliani formularon la especificación analítica que permitió la contrastación de la *hipótesis de ciclo vital*. Esta especificación fue incorporada en el modelo MPS utilizado por la Reserva Federal⁴⁰. Los trabajos derivados del citado modelo, en particular Modigliani (1971), proporcionaron las primeras estimaciones del *efecto riqueza*, y alentaron la investigación empírica⁴¹. Así comenzaron a desarrollarse trabajos basados en el modelo simple de *ciclo vital* que, mediante la aplicación de técnicas econométricas sencillas, detectaban la existencia de un cambio significativo en el consumo agregado como consecuencia de cambios previos en el nivel de la riqueza⁴².

En 1978, la introducción explícita de las expectativas en la ecuación de Euler por parte de Hall supuso la herramienta central del análisis empírico sobre el consumo en los últimos 35 años⁴³. Esta innovación fue incorporada a la posterior investigación empírica del *efecto riqueza*, lo que potenció su desarrollo.

En los años ochenta, comenzaron los primeros intentos de aplicación del análisis empírico, mediante el uso de la herramienta teórica comentada,

⁴⁰ El modelo MPS (MIT-PENN-SSRC) fue el modelo macroeconómico operativo en la Federal Reserve Board hasta 1996. Para una descripción del citado modelo, véase, por ejemplo, Rasche and Shapiro (1968).

⁴¹ A principios de la década de los setenta, Rasche (1972) y Friend and Lieberman (1975) desarrollaron los primeros estudios con encuestas familiares. No obstante, la baja calidad de los datos existentes paralizó este tipo de investigación, ante la necesidad de elaborar nuevas encuestas y datos de panel.

⁴² Véase, por ejemplo, Modigliani and Steindel (1977), Elliot (1980) y Brayton and Mauskopf (1985, 1987).

⁴³ Hall demostró que suponiendo que la teoría de la renta permanente fuera exacta y los agentes tuvieran expectativas racionales, los cambios del consumo en el tiempo serían imprevisibles. Así, el consumo actual puede expresarse como el consumo esperado más una sorpresa aleatoria e imprevisible.

alentados por el objetivo de valorar el impacto de los *crash bursátiles* de 1929 y 1987 sobre el consumo⁴⁴. Pero fue en la década siguiente, como consecuencia del rápido e inusual incremento en la riqueza de las familias que tuvo lugar en buena parte de los países desarrollados, cuando se impulsó definitivamente la investigación empírica del efecto *riqueza*.

Así, el análisis empírico del efecto *riqueza* hace preciso disponer de una estimación de referencia que permita cuantificar la cantidad por la cual una familia puede incrementar su consumo si recibe un shock de riqueza positivo. La práctica habitual en la mayoría de los estudios es utilizar el modelo postulado por Galí (1990) para estimar el efecto *riqueza*, según el cual:

$$C_t = \alpha + \beta W_t + \delta Y_t + u_t$$

En la interpretación estándar, el coeficiente β es la *propensión marginal a consumir de la riqueza* (W_t), esto es, el incremento en el gasto del consumidor asociado con un incremento unitario en la riqueza. El coeficiente δ es la *propensión marginal a consumir de la renta* (Y_t). Básicamente Galí modificó la versión de horizonte infinito de la *hipótesis de la renta permanente*, permitiendo un horizonte finito similar a la *hipótesis del ciclo vital*.

HÁBITOS DE CONSUMO

Por último, también podemos citar los hábitos de consumo como variable que influye en el propio consumo. Una manera sencilla de introducir el efecto de los hábitos de consumo consiste, como se hace por ejemplo en el modelo de Klein-Golderberg (1955), en utilizar el consumo retardado como variable explicativa. Barbancho (1961), quien utiliza este mismo argumento, llega a la siguiente estimación

$$C_t = \beta_0 + \beta_1 Y_t + \beta_2 C_{t-1} + u_t$$

siendo $\beta_0 > 0$, $0 < \beta_1 < 1$ y $|\beta_2| < 1$.

Desafortunadamente, en ninguna de las fuentes de microdatos en estudio es esta Memoria (ECV, EPF y EFF), disponemos del consumo del hogar retardado un período, por lo que no podremos utilizar esta variable en ningún caso.

2.1.2. Factores estructurales

Ahora debemos considerar los efectos sobre el consumo agregado que surgen de los cambios habidos en la importancia relativa de clases distintas de consumidores. Si hay variaciones en esta estructura de unidades de gasto, el consumo puede cambiar independientemente de las variaciones ocurridas en

⁴⁴ Conviene destacar que los principales análisis empíricos realizados en esta década concluyeron la escasa significatividad de los efectos *riqueza*. No obstante, la mayoría de los investigadores coinciden en señalar la ínfima proporción de activos bursátiles en la riqueza de las familias como principal explicación del escaso impacto de la caída bursátil de 1929 sobre el gasto del consumidor. Véase, por ejemplo, Temin (1976) y Romer (1990). Los argumentos que intentan explicar la experiencia de 1987 se basan en el carácter temporal de la caída bursátil y los efectos compensatorios de los precios alcistas en los trimestres inmediatamente anteriores y posteriores al citado *crash*. Cagan (1990) constituye un buen ejemplo en esta línea.

el ingreso agregado (o de cualquiera de las demás variables anteriormente consideradas). En particular, estudiaremos, dentro de este conjunto de efectos la propia forma en que se distribuyen la renta en el conjunto de consumidores (distribución de la renta) y los factores demográficos.

DISTRIBUCIÓN DE LA RENTA

Cuando hablamos de distribución del ingreso hacemos referencia a la forma en que se distribuye la riqueza generada en una región o en un país entre los distintos segmentos de la población que la integran en un período determinado. Así, la distribución del ingreso nos permite ubicar las condiciones de desigualdad y grados de concentración que presenta una sociedad.

Los estudios de ingreso-gasto resultan de gran importancia para observar la relación que guarda la distribución del ingreso y la estructura social del gasto en consumo en un espacio y tiempo específico. A este respecto la EPF puede ayudarnos a mostrar dicha relación para el caso de España.

Hasta la década de los 50, la teoría del consumo era dominada por la teoría de la ley psicológica fundamental aportada por Keynes (1936) que establecía que tanto la propensión marginal a consumir (PMgC) como la proporción media a consumir eran decrecientes en función del ingreso. Sin embargo, un estudio empírico realizado por Kuznets (1946) demostró que, si bien esto se cumplía para el corto plazo y para los cortes trasversales, en el largo plazo el consumo presentaba una propensión media a consumir constante que contradecía lo propuesto por Keynes.

En 1949, Duesenberry publica el libro "Income, Saving and the Theory of Consumption Behaviour" en el que realiza una crítica de las teorías del consumo vigentes hasta ese momento, por ser inconsistentes y simplistas ya que no consideraban: 1) la formación de hábitos a la hora de consumir; 2) la influencia del patrón de consumo de las personas cercanas a cada uno (lo que el autor denominó "efecto demostración"). Esto último se ve reforzado por el deseo de las personas de distinguirse del resto, que los lleva a adquirir bienes de lujo para proteger o incrementar su status social ("efecto Veblen")⁴⁵. Esto llevó a la conclusión de que es el ingreso relativo, y no el absoluto, lo que determina el consumo y el ahorro individual. Así, la crítica sociológica realizada por Duesenberry no solo apuntaba a la teoría keynesiana sino además a la teoría utilitarista ya que la aspiración de un prestigio social mayor genera la adquisición de bienes que no pueden ser considerados útiles en los términos convencionales (Mason, 2000).

Se puede observar que la inclusión de estos dos elementos (los hábitos y los efectos demostración) permitiría explicar los hechos estilizados presentados por Kuznets: si se divide la población en dos clases (pobres y ricos), al realizar un corte trasversal se observará que los pobres consumirán una mayor fracción de sus ingresos que los ricos, explicado por el efecto demostración de querer emular el consumo de estos últimos. Esto significa además que los ricos tendrán una propensión a ahorrar mayor que los pobres, y que la propensión media a ahorrar se ubicará en algún lugar entre medio de ambas. Si bien uno esperaría

⁴⁵ Por lo tanto, se puede considerar a la teoría de Duesenberry como parte de la corriente que estudia la economía con un enfoque institucionalista. En Hamilton (1987) se puede encontrar una reseña que estudia particularmente el análisis del consumo aplicado por esta corriente.

ver que al extender este análisis al largo plazo la propensión media a ahorrar aumente (porque se incrementa el ingreso de toda la población), esto no sucede debido a que lo que importa es el ingreso relativo y no el absoluto. Aun cuando el ingreso de toda la población aumente, siempre existirá una parte de la población relativamente más pobre que la otra, lo que implica una propensión a ahorrar también menor que la otra, que en el agregado se mantiene constante. Por otro lado, los hábitos en el consumo permitirían explicar el comportamiento más suave que presenta el consumo en relación a las variaciones de los ingresos en el corto plazo, principalmente ante caídas bruscas observadas en las recesiones.

Por tanto, en base a esto y tal y como corroboran estudios como los de Carrol (2000), se puede suponer que (puesto que se cumple que $1 = PMgC + PMgS$), la propensión marginal a consumir de las familias con ingresos bajos es sustancialmente más alta que la de las familias de ingresos altos.

FACTORES DEMOGRÁFICOS

Cuando examinamos un estudio de sectores representativos hallamos que las diferencias en el consumo aparecen mostradas claramente por niveles de ingreso. Pero aun a cualquier nivel de ingresos dado, hay amplias diferencias entre el gasto de consumo de las familias individuales. Estas diferencias pueden ser explicadas por lo menos en parte por factores demográficos. Los estudiamos a continuación.

Tamaño de la familia

Un factor que actúa sobre el consumo es el tamaño de la familia. Esta variable desempeña un papel de primer orden cuando se trabaja con datos espaciales o atemporales (*cross-section*), como es nuestro caso. En caso de series temporales su importancia es muy secundaria, ya que el tamaño de las familias es una variable de gran estabilidad en el tiempo.

Existen multitud de estudios que demuestran que, *ceteris paribus*, el gasto de las familias grandes es mayor que el de las más pequeñas. Así, por ejemplo, en Kiran y Dawan (2015) se confirma que el tamaño de la familia tiene un impacto positivo significativo en el gasto de consumo de los trabajadores industriales. Además, la reducción del ahorro debido a la mayor carga del miembro adicional de la familia corrobora el efecto negativo propuesto del tamaño de la familia en el ahorro de los trabajadores.

El "ciclo vital" de la familia

En la práctica se sabe que las familias con niños muy jóvenes -o hijos en la universidad- tienden a gastar más que otras familias del mismo tamaño. Es decir, el ciclo de vida de la familia en que se encuentra la persona también tendrá una influencia sobre sus valores, estilos de vida y comportamiento de compra en función de si es soltero/a, parejas casadas, parejas de hecho, tiene una relación, tiene niños, etc. Asimismo, un consumidor no compra los mismos productos o servicios a los 20 ó 70 años. Su estilo de vida, valores, medio ambiente, actividades, aficiones y hábitos de consumo evolucionan a lo largo de su vida. La edad y el ciclo de vida tienen un impacto potencial sobre el

comportamiento de compra del consumidor. Es obvio que los consumidores cambian la compra de bienes y servicios con el paso del tiempo.

En particular, son numerosas las referencias que, de un modo u otro, basándose por lo general en argumentos de la *teoría del ciclo vital*, han contrastado la significatividad de la edad de los individuos sobre el consumo privado en España. En su mayor parte, se trata de trabajos que utilizan observaciones individuales procedentes de la Encuesta de Presupuestos Familiares del INE, y en los que la “edad de las familias” se aproxima por medio de la edad del cabeza de familia o sustentador principal. Tal es el caso de García y Labeaga (1996), López-Sálido (1993), y Labeaga y López (1996) en donde se estima un sistema completo de ecuaciones de demanda (o, en su defecto, la demanda de un determinado bien o servicio), y de Cortázar (1996) en donde se estima el gasto total. Deaton, Ruiz-Castillo y Thomas (1989), Labeaga y López (1994) también estiman un sistema de demanda, y López (1993) estima la demanda de un bien particular, pero considerando la distribución por grupos de edades de la unidad familiar. En todos los casos, la variable edad del cabeza de familia (y, en su caso, también el cuadrado de ésta) y las proporciones de los diferentes grupos de edad resultan significativas.

En el panorama internacional podemos destacar trabajos como el de Gourinchas & Parker (2002) demuestran que, utilizando datos a nivel individual para construir perfiles promedio de los ingresos y el consumo durante la vida laboral de los hogares, el consumo mantiene la forma de joroba, incluso después de controlar los efectos de la familia y de cohorte.

Lugar de residencia

La región del país y el tipo de ciudad en la que vive un individuo (ciudad grande, pequeña ciudad, campo, etc.) también tendrán una influencia sobre sus valores, estilos de vida y comportamiento de compra. Por ejemplo, es más que posible que los consumidores que viven en Madrid no tengan el mismo comportamiento y hábitos de compra que los de Badajoz.

Numerosos estudios demuestran la existencia de las relaciones entre el consumo y el lugar de residencia. Es el caso de Gracia *et al.* (1998), que estudia las diferencias existentes entre consumo rural y urbano, estableciendo el límite de separación entre ambos entornos, siguiendo la metodología empleada por diversas instituciones, en los municipios de 10.000 habitantes. Según este estudio, efectivamente, el comportamiento del consumidor depende del tamaño del municipio en el que habita. En concreto, se atribuye dicho comportamiento diferencial a la existencia de precios relativos significativamente diferentes en los distintos municipios y, en menor medida, al nivel de consumo existente. Por su parte, Moreno (2006) considera que el lugar de residencia fue un factor decisivo en la configuración de las pautas de consumo de la población catalana del siglo XVIII, si bien el grado de influencia de la variable residencial no era igual para todos los grupos sociales y no tenía las mismas consecuencias en el modo de vida.

Ocupación

Respecto a la ocupación, parece ser también una variable candidata a relacionarse con el consumo. Trabajos como el de Cage (1989) ponen de manifiesto que existen diferencias entre los distintos grupos ocupacionales con

respecto a sus distribuciones de gastos. Asimismo, el análisis demostró que la variable ingreso, determinada en mayoría de los casos por la ocupación, fue una fuerte impulsora en la determinación del nivel de gasto para todas las categorías de gastos investigados.

Nivel de educación

La forma más habitual de medir los beneficios netos de la educación es a través de la función de ganancias de Mincer, utilizando el llamado modelo de escolaridad, en el que se supone que todo el capital humano se ha adquirido exclusivamente durante el periodo de formación escolar.

Mincer (1974) supone que, si el único coste de un año más de educación es el coste de oportunidad del tiempo empleado por los estudiantes y, si el aumento proporcional en los ingresos que se obtiene por la educación adicional recibida es constante a lo largo de la vida, entonces el logaritmo de los ingresos estará relacionado de forma lineal con los años de educación de las personas y la pendiente de esta relación puede interpretarse como la tasa de rendimiento de la inversión en educación. Mincer añadió a su modelo un término para tener en cuenta los rendimientos producidos por la formación recibida en el puesto de trabajo. En concreto:

$$\ln W_i = \beta_0 + \beta_1 S_i + \beta_2 X_i + \beta_3 X_i^2 + \varepsilon_i$$

siendo $\ln W_i$ el logaritmo natural del salario de la persona i , S_i los años de escolaridad, X_i la experiencia, y X_i^2 la experiencia al cuadrado (que mide si el rendimiento de la misma es decreciente, constante o creciente). ε_i es un término de perturbación aleatorio que recoge las variaciones salariales no capturadas por los regresores incluidos en el modelo. Y el coeficiente β_1 sería la tasa de descuento o rendimiento de un año adicional de educación, que pretende recoger la prima salarial a la educación; esto es, mide el incremento porcentual de los salarios atribuible a un aumento de un año en el nivel de escolarización, controlando por la experiencia laboral y suponiendo que es independiente del nivel educativo.

Existen otros muchos trabajos de los que se desprende la relación entre la educación y la capacidad para ganar en el mercado laboral. Numerosos estudios en los últimos diez años han demostrado que las ganancias se elevan de forma sistemática por la educación y que la compra de un cierto nivel de educación superior es una sabia inversión para la mayoría de la gente cuando el retorno de la inversión se mide en resultados mejorados por sí solos.

Por otro lado, Michael (1975) se centra en identificar y describir el efecto de la educación en otro aspecto de la conducta: los patrones de gasto de los consumidores. En concreto, estima que un incremento del 10% en el nivel de estudios (por ejemplo, respecto a la media de 10 años a 11 años) es equivalente a elevar el nivel de los gastos totales de la unidad familiar en un 0,8%.

Según Beaton (1975) las personas con más educación, en promedio, reciben salarios más elevados que las de la misma capacidad, pero con menos educación. Pero sin duda para este autor el grupo más llamativo en el estudio son los trabajadores por cuenta propia, ya que sus salarios son

sustancialmente más altos que los salarios de los demás con la aptitud y la educación equivalente.

Finalmente, podemos encontrar trabajos como los de Cristofani (2011); "Los determinantes para el ahorro en los hogares de bajos ingresos del GBA según los enfoques teóricos", que realiza sus estimaciones con datos de cortes transversales, (los datos son de la encuesta nacional de gastos de hogares 1996-97) poniendo su atención en los hogares de bajos ingresos. En él se encuentran relaciones positivas con los niveles de ingresos y niveles de educación (a mayor cantidad de ingresos y nivel de educación, mayor capacidad de ahorro).

Género

La gran mayoría de los estudios de las desigualdades en el consumo se han caracterizado por el uso de la clase social como el principal concepto para explicar estas desigualdades. Sólo unos pocos estudios introducen el concepto de género en la explicación del comportamiento de los consumidores.

Desde la economía y también desde la sociología se ha considerado al hogar como la unidad básica de consumo, tan solo hay que pensar en las bases de datos oficiales que existen, como es la Encuesta Continua de Presupuestos Familiares, para corroborar esta idea. Ello ha implicado que sean pocos los estudios y los datos cuantitativos que se dispongan para analizar las diferencias y desigualdades existentes entre hombres y mujeres respecto al consumo.

Borrás (2007) aporta algunas reflexiones teóricas que contribuyen al estudio de las desigualdades en el consumo entre los géneros. También aporta algunas evidencias empíricas que nos muestran cómo el género es un eje explicativo de las desigualdades en el consumo. Según este estudio hay que considerar varios aspectos para estudiar las desigualdades del consumo por razón de género. En primer lugar, la capacidad para decidir el tipo de consumo, es decir quién decide en qué y cuánto se gasta. En segundo lugar, quién realiza el acto de compra, y ligado a la misma distinguir entre compra placer y compra trabajo. Y por último el beneficiario o beneficiarios de dicha compra. Teniendo todo esto en cuenta, se llega a que las compras que se realizan con mayor frecuencia y son necesarias para el bienestar del hogar forman parte del trabajo cotidiano de las mujeres, mientras que los hombres participan de aquellas que implican mayor gasto y están cargadas de representación social, donde el cabeza de familia ha de estar, por supuesto, presente.

Por su parte, para Lenormand *et al.* (2015), las mujeres tienden a gastar más dinero en ropa, comida, hipermercados, salud y bienestar y más cerca de su casa que los hombres, que concentran su gasto en coches, bares, restaurantes, tecnología y transporte. Así lo demuestra el estudio "*Influence of sociodemographic characteristics on human mobility*" realizado por los investigadores del Instituto de Física Interdisciplinar y Sistemas Complejos, IFISC (CSIC-UIB) sobre 40 millones de transacciones comerciales realizadas con tarjetas bancarias BBVA en las ciudades españolas más pobladas, Barcelona y Madrid durante el año 2011. Tanto en Madrid como en Barcelona los resultados fueron similares. En ambos casos, el estudio geolocalizado de las

transacciones con tarjetas bancarias demostró que los hábitos de consumo y patrones de movilidad varían en función del género, la edad y la ocupación.

Propiedad de viviendas

La vivienda es un componente importante de la riqueza. Dado que los precios de las casas fluctúan considerablemente con el tiempo, es importante entender cómo estas fluctuaciones afectan a las decisiones de consumo de los hogares. El aumento del precio de la vivienda puede estimular el consumo mediante el aumento de la riqueza percibida de los hogares, o mediante la relajación de las restricciones de crédito. Campbell & Cocco (2004) investigan la respuesta del consumo de los hogares a precios de la vivienda utilizando microdatos del Reino Unido. Según estos autores, los efectos más grandes de los precios de la vivienda sobre el consumo recaen en los propietarios de edad avanzada, y el pequeño efecto, sin importancia, para los inquilinos más jóvenes. Este hallazgo es consistente con la heterogeneidad en los efectos riqueza a través de estos grupos. Se sugiere que a medida que la población envejece y se concentra más en el grupo de propietarios de viviendas antiguas, el consumo agregado puede ser más sensible a los precios de las casas.

Además, para Campbell & Cocco el consumo responde a los cambios previsibles en el precio de la vivienda, lo cual no sólo es consistente con un aumento de los precios inmobiliarios en un contexto de restricciones de crédito, sino que también se puede explicar por la causalidad inversa de la liberalización financiera en los precios de las casas, o por el ahorro motivo precaución. Los efectos sobre el consumo de los cambios previsibles en los precios de las casas parecen ser más débiles para las familias que tienen capacidad de endeudamiento no utilizado, pero afectan a inquilinos y propietarios de viviendas y trabajan a través de los precios nacionales de las casas en lugar de los regionales. Esto sugiere que los precios de la vivienda del Reino Unido están relacionados con la facilidad o dificultad de pedir préstamos en la economía en su conjunto.

Por su parte, Cristofani (2011) define el Ahorro como Activo (motivo precaución) como:

$$\text{Ahorro (precaución)}_t = \beta_0 + \beta_1 \text{Ingreso}_t + \beta_2 \text{Régimen Tenencia del hogar}_t + \beta_3 \text{Deuda por Adquisición de vivienda}_t + u_t$$

Según este estudio, la tenencia de la propiedad contribuye de manera positiva en el ahorro. En cambio, la marca de deuda por adquisición del bien tiene un efecto positivo sobre el ahorro, aunque el coeficiente estimado no es significativamente distinto de cero.

Raza

Finalmente, dentro del conjunto de factores demográficos determinantes del consumo estudiaremos la influencia de la raza de los consumidores. Tradicionalmente, en países donde existe una importante diversidad de razas, como Estados Unidos, se ha comprobado que los hogares negros tienden a tener niveles más bajos de gasto total que los blancos (y, por ende, mayor ahorro) y que estas diferencias tienden a persistir en todos los niveles de ingresos. Así lo constatan trabajos como los de Charron-Chénier *et al.* (2017),

quienes, utilizando los datos de las encuestas de gasto, examinan el consumo desde una perspectiva estructural y argumentan que los hogares negros se enfrentan a restricciones únicas que restringen su capacidad para adquirir bienes y servicios importantes. En general, estos análisis indican que existen diferencias raciales en el consumo independientemente de otras disparidades económicas, por lo que puede considerarse un factor de interés en el estudio de la desigualdad.

Una vez estudiados los factores que influyen en el consumo de los hogares y, por tanto, en el ahorro, disponemos de una amplia variedad de relaciones teóricas que, partiendo de nuestra base de datos de hogares que contiene información sobre los ingresos y características sociodemográficas de éstos declarados en la ECV-la que hemos denominado BDH-, utilizaremos en el siguiente apartado para incorporar información complementaria procedente de otras fuentes microdatos (EPF, para el estudio del consumo, y EFF, para la situación financiera de los hogares), por medio de técnicas de fusión estadística, con la finalidad de ampliar la información contenida en nuestra BDH y aflorar, por medio de dichas relaciones, parte de la sub-declaración de ingresos presente en la ECV.

3. Metodología propuesta para estimar el consumo y la situación financiera de los hogares de la ECV

La constatación en el capítulo II de que la renta disponible de los hogares en la ECV está infravalorada respecto a los resultados publicados por la CNE y de que presenta ciertas diferencias en la definición de sus componentes, careciendo incluso de algunos de ellos, justifica la búsqueda de algún procedimiento de mejora de la calidad de la misma. Así pues, el objetivo de este apartado es presentar de forma detallada el método que se empleará para la estimación de los gastos y la situación financiera de los hogares de la ECV, con la finalidad de mejorar la consistencia global de la información contenida en esta última.

Como ya venimos comentando, las encuestas a hogares especializadas en recopilar esta información en España son: la Encuesta de Presupuestos Familiares (EPF), que suministra amplia información de los gastos de los hogares junto a otras menos detallada de sus ingresos; y la Encuesta Financiera de las Familias (EFF), que proporciona información detallada de las rentas, los activos, las deudas y el gasto de cada unidad familiar. Asimismo, sabemos que la ECV es la fuente que recoge los ingresos de los hogares con el mayor nivel de detalle, y sobre la cual deseamos estimar variables económicas presentes en las anteriores encuestas. Sin embargo, el hecho de que dicha información sobre ingresos, gastos y situación financiera se recoja básicamente de forma separada en estas tres encuestas y sobre hogares diferentes, dificulta la posibilidad de efectuar una comparación y, en consecuencia, la realización de aquella plausible corrección de los datos de ingresos en la ECV.

Teniendo en cuenta el mayor grado de detalle que proporcionan las ECV sobre los ingresos de los hogares, en este trabajo se propone la utilización de

técnicas de fusión estadística, en una primera fase, entre la base datos que hemos construido con la ECV -BDH- y la EPF y, posteriormente, entre el resultado de aquella y la EFF, para los años de nuestro estudio, 2008 y 2011. El resultado que obtengamos de este ejercicio será un nuevo conjunto de datos más completo y robusto, que denominamos muestra de fusión o muestra *match*, cuya riqueza de información nos permitirá identificar, particularmente en cada tipo de hogar: por un lado, la relación existente entre los niveles de gastos de la EPF e ingresos declarados en la ECV y, por otro, entre el nivel de riqueza y la situación financiera de las familias en la EFF y aquellos ingresos, con el fin de proceder a corregir las rentas declaradas en la ECV con base en aquellas relaciones y entre las que hemos estudiado en el anterior apartado de este capítulo.

Por todo lo comentado con anterioridad, a lo largo de este apartado llevaremos a cabo el estudio del método de fusión estadística que utilizaremos para imputar a los hogares de la ECV valores de gasto de la EPF y variables relacionadas con la situación financiera de éstos a través de la EFF. Así, en primer lugar, estudiaremos la comparabilidad y homogeneidad de las variables de interés proporcionadas por dichas fuentes, como paso necesario y previo para iniciar el proceso de fusión estadística de éstas. Posteriormente, describiremos el proceso de fusión estadística empleado en este trabajo para capturar la relación, en primer lugar, entre los niveles de los ingresos y los gastos declarados por los hogares y, después, entre los primeros y la riqueza de las familias. Para ello, nos detendremos en las descripciones de la medida empleada para cuantificar la proximidad entre los hogares de ambas fuentes, del método de enlace empleado para casar los casos fusionados, del proceso de selección y de ponderación de las variables comunes empleadas para la fusión, junto a la especificación de la información a fusionar.

3.1. La Fusión Estadística de la BDH (ECV), la EPF y la EFF

3.1.1. Fundamentación teórica

De forma general, podemos decir que los procedimientos estadísticos de fusión (*statistical matching*), constituyen un método de imputación y/o estimación múltiple de variables observados en diferentes fuentes estadística, cuyas unidades de análisis generalmente provienen de una misma población y poseen variables en común pero no se superponen (Leulescu y Agafitei, 2013). Los posibles beneficios de este enfoque radican en la posibilidad de mejorar el uso complementario y el potencial analítico de fuentes de datos existentes (por ejemplo, información estadística transversal que abarca una amplia gama de aspectos socioeconómicos, como es nuestro caso). Por lo tanto, la fusión estadística puede ser utilizada como una herramienta para aumentar la eficiencia de uso de los conjuntos de datos existentes con el fin de estudiar la relación existente entre algunas variables no observadas conjuntamente.

El objetivo principal del *statistical matching* (SM) se puede lograr mediante dos enfoques diferentes según el resultado que se desee. D'Orazio *et al.* (2006) los define como enfoques micro y macro:

- o En la aproximación micro, que será la que utilizaremos en nuestro trabajo, el objetivo es generar un fichero sintético con información completa, transformando los distintos conjuntos de datos en uno solo integrado cuyos registros se refieren a la misma unidad de análisis y donde todas las variables de interés están presentes. Este fichero sintético posteriormente se emplea para realizar análisis conjuntos referentes a variables que inicialmente se encuentran en ficheros separados. El término "sintético" se refiere al hecho que este archivo no es el resultado de observaciones directas de todas las variables sobre un conjunto de individuos que pertenecen a la población de interés.
- o El objetivo del enfoque macro es transformar los conjuntos de datos en resultados agregados, identificando estructuras que describan la relación entre las variables de las dos fuentes que no han sido observadas de manera conjunta (tablas de contingencia, matrices de correlación, distribuciones conjuntas y marginales, etc.).

La correspondencia entre las bases de datos que intervienen en una fusión se establece a partir de del emparejamiento de observaciones que se consideran semejantes, teniendo en cuenta algún criterio de proximidad o distancia. Es necesario que exista, entre las fuentes de información a fusionar, una zona de intersección formada por variables comunes a las que denominamos *variables de fusión* o *variables match* y las representamos por el vector $X(.)$.

Siguiendo a Calonge y Manresa (2001), el caso típico del SM de dos fuentes de datos se puede explicar del siguiente modo: Dados dos conjuntos de datos A y B provenientes de dos encuestas muestrales independientes con información sobre variables (Y, X) y (X, Z) respectivamente, donde X son las variables comunes observadas en ambas fuentes es decir, que tienen la misma o similar definición, y correlacionadas con Y y Z; Y representa a las variables específicas observadas solamente en A; y Z a las variables específicas observadas sólo en B, de manera que Y y Z no se observan de manera conjunta (Kadane, 1978). La relación entre las variables comunes y las variables específicas observadas sólo en uno de los conjuntos de datos -el conjunto de datos del donante- se explorará y se utilizará para imputar a las unidades del otro conjunto de datos -el conjunto de datos receptor- las variables no observadas directamente. Así, se genera el resultado final que denominaremos muestra fusionada o muestra *match*, que incluye el siguiente vector de variables: $(Y_{BASE}, X_{BASE}, X_{NOBASE}, Z_{NOBASE})$, donde Y_{BASE} y Z_{NOBASE} representan las variables no comunes del conjunto de datos receptor y donante, respectivamente. La elección de una de las muestras como base depende del uso y las aplicaciones a realizar con la nueva información. En general, la muestra base es la fuente de información estadística que consideramos más importante. La muestra no-base actúa entonces, utilizando la terminología del diseño de experimentos, como *grupo de control*, y su información es asignada a la muestra base dando lugar a la muestra de fusión (cuadro 3.3).

Cuadro 3.3
ESTRUCTURA DE LA FUSIÓN ESTADÍSTICA

<i>Muestra Base</i>	1 . . . n_{BASE}	Y_{BASE}	X_{BASE}	
<i>Muestra No Base</i>	1 . . . $n_{NO\ BASE}$		$X_{NO\ BASE}$	$Z_{NO\ BASE}$

→

<i>Muestra Match</i>	1 . . . n_{BASE}	Y_{BASE}	X_{BASE}	
		$X_{NO\ BASE}$	$Z_{NO\ BASE}$	

Fuente: Calonge y Manresa, (2001).

Sin embargo, este punto de vista es bastante simplista y se ha planteado una preocupación metodológica importante con respecto a la validez de los resultados ya que en cada caso particular el investigador se enfrenta a un gran número de problemas ad hoc (Sutherland *et al.*, 2001). Los orígenes⁴⁶ de la fusión estadística se remontan a mediados de los años sesenta, cuando en Estados Unidos se compararon los archivos de impuestos de 1966 y la Encuesta de oportunidades económicas de 1967 con el fin de proporcionar un conjunto de datos sintéticos sobre variables sociodemográficas. A principios de la década de 1970, se aplicaron diferentes técnicas de adaptación a las encuestas sociales en los Estados Unidos (Ruggles and Ruggles, 1974), pero estas técnicas fueron severamente criticadas basándose en suposiciones no justificadas ni verificables (Kadane, 1978; Rodgers, 1984). En particular, las medidas de asociación entre Y y Z condicionales a X no se pueden estimar y se asume generalmente que son 0. Se trata de la denominada *suposición de independencia condicional*⁴⁷ (CIA), un punto de referencia para evaluar la calidad de las estimaciones basadas en técnicas de SM.

⁴⁶ En España la técnica de fusión estadística ha sido utilizada en los trabajos de Alegre *et al.* (2001) y Calonge y Manresa (2001) con el objetivo de disponer en una misma muestra de datos de gasto y de ingreso de las familias, estos autores realizaron una fusión entre Panel y la Encuesta de Presupuestos Familiares de 1990. Por su parte, en Picos (2003), con el objetivo dar una solución a los problemas de infravaloración y escasa desagregación que presentan los datos de rentas de capital que se recogen en el Panel de Hogares de la Unión Europea (PHOGUE), se imputan a PHOGUE, mediante técnicas de fusión estadística, valores de rentas de capital provenientes del Panel de Declarantes de IRPF del Instituto de Estudios Fiscales, con lo que se consigue mejorar la calidad de las simulaciones que se lleven a cabo con estos datos a la hora del estudio de la implementación de políticas sociales, como pueden ser la relativas a la concesión de pensiones de jubilación.

⁴⁷ Cuando el SM se basa sólo en las variables X comunes (o en un subconjunto adecuado), el analista está imponiendo un modelo probabilístico específico entre las variables X, Y y Z: la independencia condicional de Y y Z dada X (*hipótesis de Independencia Condicional* -CIA-). Dos variables aleatorias Y y Z son condicionalmente independientes dado una tercera variable aleatoria X si y sólo si son independientes en su distribución de probabilidad condicional dada X. Esto es, Y y Z son condicionalmente independientes dado X si y sólo si, dado cualquier valor de X, la distribución de probabilidad de Y es la misma para todos los valores de Z y la distribución de probabilidad de Z es la misma para todos los valores de Y. Esta suposición implica que las variables X explican completamente la relación entre Y y Z.

Cuando esta condición se mantiene, los algoritmos coincidentes producirán estimaciones precisas que reflejen la verdadera distribución conjunta de las variables que fueron recolectadas en múltiples fuentes. Proporcionará los mismos resultados que un procedimiento de enlace perfecto. En caso de que la independencia condicional no se mantenga, y no haya información adicional disponible, el modelo tendrá problemas de identificación y los conjuntos de datos artificiales producidos pueden conducir a inferencias incorrectas. En nuestro caso, y según se ha visto en las teorías económicas anteriores de este capítulo, las relaciones más directas, tanto para los gastos como para la riqueza que serán las variables que añadiríamos a las de la BDH mediante los procedimientos de fusión, se dan con los ingresos y demás factores objetivos y estructurales que controlaremos como variables comunes en nuestros procedimientos de *matching*. Hecho esto, parece razonable asumir la hipótesis de independencia condicional, y esperar resultados razonablemente aceptables.

3.1.2. Fases de la Fusión Estadística de la BDH (ECV), la EPF y la EFF

El procedimiento de fusión abarca una serie de etapas previas que según nuestra experiencia y la diversidad de aplicaciones que existen en la literatura, podemos resumir en las siguientes:

- a) Definición del papel que tendrá cada muestra en el proceso de fusión e identificación en éstas de las variables potencialmente comunes y no comunes.
- b) Armonización y conciliación de fuentes: homogenización de las muestras y transformación de variables.
- c) Selección de las variables de fusión.
- d) Elección del tipo de fusión y del criterio de función de distancia a aplicar.
- e) Aplicación y evaluación de resultados.

A continuación, realizamos una breve descripción de cada una de estas etapas en la fusión llevada a cabo entre la ECV y la EPF, en primer lugar, y entre el resultado de las anteriores y la EFF.

3.1.2.1. Definición del papel de cada muestra e identificación de variables

La primera decisión que debe tomarse en relación con la fusión es el papel que va a desempeñar cada una de las bases de datos utilizadas. Lo razonable en nuestro caso será utilizar la ECV como muestra principal, siendo la EPF y la EFF muestras complementarias o donantes. De esta forma, la estructura y el tamaño de la muestra fusionada serán los de la ECV, y el papel de la EPF y de la EFF consistirá exclusivamente en proporcionar a cada observación de la ECV información sobre los gastos y la situación financiera, respectivamente. Además, es necesario que las muestras presenten un cierto número de

variables comunes, proporcionando además cada una de ellas variables propias no comunes, cuestiones ambas que se cumplen en nuestro caso.

3.1.2.2. Armonización y conciliación de fuentes

Población de referencia, ámbito geográfico y temporal y unidades de análisis

En la práctica, la información de que disponemos puede requerir algún tipo de tratamiento con la finalidad de homogeneizar las muestras respecto al mismo universo de referencia. En general, los objetivos de la homogeneización son conseguir que las unidades de observación de cada una de las dos muestras sean semejantes, las poblaciones de referencia y ámbito geográfico de las mismas sean compatibles y las variables comunes tengan una definición idéntica o muy similar.

En cuanto a las unidades estadísticas para las que se proporciona información, si bien se tratan tanto de individuos como de hogares en las encuestas que queremos fusionar (ECV, EPF y EFF), debido a que el objetivo perseguido en este trabajo sobre el estudio de la sub-declaración de ingresos en la ECV está centrado sobre los hogares, en el procedimiento de fusión que llevaremos a cabo tomaremos el hogar como sujeto de análisis.

Por su parte, la población, ámbito geográfico y temporal de referencia que abarcan las encuestas que pretendemos fusionar en este trabajo (ECV, EPF y EFF) es la población residente en España que vive en hogares privados en el año natural anterior a la realización de la encuesta en el primer caso, y en el mismo año de ejecución de la misma en el caso de la EPF y EFF. La misma vivienda, el hecho de compartir los recursos económicos, el mantenimiento común y los lazos familiares son los criterios principales y más utilizados para identificar un hogar. Entre la ECV y la EPF no se aprecian apenas diferencias, sin embargo, sí se detectan algunas distinciones metodológicas entre las anteriores y la EFF, en términos de: (a) la duración del período de ausencia y (b) el tratamiento del servicio doméstico.

Por ejemplo, en la ECV y en la EPF se consideran hogares ausentes temporalmente (hospital, clínica u otra situación) y, por tanto, no serán miembros de éste, aquellos cuya duración real o prevista de su ausencia del hogar es mayor o igual a seis meses. En cuanto al servicio doméstico, ambas encuestas matizan que se considerará parte del hogar si comparten los gastos del mismo y en la actualidad no tienen otro domicilio privado que consideren su residencia habitual. En la EFF se considerará que una persona forma parte del hogar si ha estado presente en el mismo durante al menos tres meses, mientras que no incluye el servicio doméstico en ninguno de los casos.

El análisis preliminar de los datos a nivel nacional parece indicar que estas diferencias no tienen efectos significativos sobre la comparabilidad y, por lo tanto, podemos considerar que las tres poblaciones se superponen en gran medida.

Homogeneización de variables comunes

En este apartado, iniciamos la búsqueda, en el fichero receptor o conjunto de datos base (BDH base 2013, que contiene los datos de la ECV en dicha base), de variables homogéneas con algunas de las que se recogen en los

ficheros donantes o conjuntos de datos no base (EPF y EFF), con las que posteriormente valoraremos su posible utilización como variables comunes de control (variables de fusión) para la fusión estadística de dichas fuentes.

Dado que el objetivo que pretendemos conseguir mediante este proceso de fusión consiste en mejorar la estimación de los ingresos de los hogares de la ECV –recogidos en la BDH base 2013- con la ayuda de la información auxiliar presente en la EPF y EFF, y sabiendo que el ingreso familiar es un flujo que permite consumo y contribuye a los cambios en la riqueza o del patrimonio neto del hogar -cuyas relaciones extraídas de la Teoría Económica hemos estudiado en el apartado 2.1 de este capítulo-, vamos a servirnos de aquellas relaciones, junto a otras circunstanciales relacionadas con los ingresos, para la elaboración de dicho listado de variables comunes.

Estudiando de forma exhaustiva los ficheros que contienen información sobre los hogares en las encuestas mencionadas (Fichero de datos básicos del hogar, Fichero de datos detallados del hogar, Fichero de datos básicos de la persona y Fichero de datos detallados de los adultos, en el caso de la ECV; Fichero de hogar, Fichero de los miembros del hogar y Fichero de gastos, en el caso de la EPF; y Ficheros de datos básicos de los cuestionarios -sección6 y otras_secciones- y Ficheros *databol* de resultados en la EFF), los cuales ponen los organismos promotores de dichas encuestas a disposición de los usuarios, se han extraído una serie de variables finalmente comunes en ambas encuestas.

En este proceso de comparación nos hemos encontrado que, a veces, en sendas fuentes existían distintos números de categorías para variables inicialmente análogas, por lo que hemos procedido a su homogeneización con el fin de que su definición sea idéntica en ambas encuestas. En estos casos, la correspondencia entre las categorías de estas variables comunes ha sido, en general, bastante sencilla de realizar, a la vez que ha sido también suficiente con la agregación de categorías de alguna de las variables para conseguir la homogeneización de las mismas en ambas encuestas.

Sin embargo, en el proceso de construcción de las variables homogéneas para la ECV, la EPF y la EFF, han surgido algunas dificultades que hemos tenido que solventar sobre los ficheros de microdatos de la ECV, EPF y EFF por medio de sendos programas de homogeneización elaborados mediante el programa estadístico SPSS y que se adjuntan en el Anexo 1 de este trabajo.

Así, por ejemplo, en el caso de la variable *Tipo de hogar*, hemos tenido que construir en la EPF y en la EFF una variable idéntica a la de la ECV (que hemos denominado TIPHO GAR_m) con la información presente en sendos ficheros de microdatos, ya que ninguna de las definiciones de las variables relacionadas con esta característica, existentes en los ficheros de usuario, era directamente comparable con la de la ECV. En cuanto a la variable *Número de miembros dependientes* económicamente (N_DEP) atendiendo a la definición que da la ECV, su elaboración en las tres encuestas ha sido posible a partir de: la creación de una variable binaria auxiliar (DEP) que toma el valor uno en caso de que un determinado miembro del hogar tenga una edad inferior a dieciocho años o bien tenga una edad inferior a veinticinco años y su situación en relación con la actividad económica sea la de inactivo, y cero en caso de que no se cumpla ninguna de las premisas anteriores.

En los cuadros 3.4 y 3.5 expresamos de forma sintética, para la BDH base 2013 y la EPF, y para la BDH base 2013-EPF (BDH base 2013 resultante tras su fusión con la EPF) y la EFF, respectivamente, la denominación y descripción de cada una de las variables comunes que se han obtenido fruto de este proceso de homogeneización. Como se puede observar en ella, coexisten dos tipos de variables: las relativas al hogar en su conjunto y aquellas cuya información está referida a alguno de los miembros del hogar (*sustentador principal*⁴⁸, cuyas variables respectivas se han nombrado, a partir de la nomenclatura común, seguidas del sufijo '_SP', y/o *segundo adulto en importancia económica*⁴⁹, para los que sus respectivas variables se han nombrado análogamente, seguidas del sufijo '_2A').

⁴⁸ El *sustentador principal* se define como aquel miembro del hogar de 16 o más años cuya aportación periódica, no ocasional, al presupuesto común se destina a atender los gastos del hogar en mayor grado que las aportaciones de cada uno de los restantes miembros.

⁴⁹ El *segundo adulto en importancia económica* es el que más ingresos aporta al hogar después del sustentador principal o, en caso de tener el mismo nivel de ingresos, el de mayor edad.

Cuadro 3.4 Variables homogéneas en la BDH base 2013 y en la EPF (*)

VARIABLES HOMOGÉNEAS RELATIVAS AL HOGAR			
VARIABLE	DESCRIPCIÓN	VARIABLE	DESCRIPCIÓN
Pais	País	INTOTALES_m	Ingresos totales netos del Hogar
ano	Año de referencia de los ingresos	IIC10_INTOTALES_m	Intervalo Interdecílico de los Ingresos totales netos del Hogar
idh	Identificador del Hogar	PI_INTOTALES_m	Percentil de los Ingresos totales netos del Hogar
w	Factor de elevación del Hogar	LINTOTALES_m	Logaritmo de los Ingresos totales netos del Hogar
nmh	Nº de miembros del Hogar	IIC10_lintotales_m	Intervalo Interdecílico del Logaritmo de los Ingresos totales netos del Hogar
nmh16	Nº de miembros del Hogar con 16 años o más	ININTOTALES_m	Intervalo de Ingresos totales netos del Hogar
nmh14	Nº de miembros del Hogar con 14 años o más	IND0_IT_m	Índice 0 del Total de Ingresos Netos del Hogar
REGION_m	Región de ubicación del Hogar	IND1_IT_m	Índice 1 del Total de Ingresos Netos del Hogar
HAB_m	Habitat del Hogar	IIC10_IND1_IT_m	Intervalo Interdecílico del Índice 1 del Total de Ingresos Netos del Hogar
TIPOVIV_m	Tipo de vivienda	I_CPRO	Indicador sobre si el Hogar tiene Ingresos por Cuenta Propia
REGTENEVIV_m	Régimen de tenencia de la vivienda	I_CAJE	Indicador sobre si el Hogar tiene Ingresos por Cuenta Ajena
NUMHABIT_m	Nº de habitaciones de la vivienda	I_PENS	Indicador sobre si el Hogar tiene Ingresos por Pensiones Contributivas y No Contributivas
TIPHOGAR_m	Tipo de Hogar	I_DES	Indicador sobre si el Hogar tiene Ingresos por Desempleo
n_dep	Nº de dependientes del Hogar (según la ECV)	I_OSPS	Indicador sobre si el Hogar tiene Ingresos por Otras Subvenciones y Prestaciones Sociales
n_percep	Nº de perceptores del Hogar	I_RCYP	Indicador sobre si el Hogar tiene Ingresos por Rentas del Capital y la Propiedad
n_hombre	Nº de hombres en el Hogar	I_OTRI	Indicador sobre si el Hogar tiene Ingresos por Otros Ingresos
n_mujer	Nº de mujeres en el Hogar	FP_ING_m	Fuente Principal de los ingresos regulares en el Hogar (clasif.1)
n_soltero	Nº de solteros en el Hogar (de 16 ó más años)	FP_ing2_m	Fuente Principal de los ingresos regulares en el Hogar (clasif.2)
n_casado	Nº de casados en el Hogar (de 16 ó más años)	VARIABLES HOMOGÉNEAS RELATIVAS A LOS MIEMBROS DEL HOGAR (i=SP -Sustentador Principal-, 2A -Segundo Adulto en importancia económica-)	
n_separado	Nº de separados en el Hogar (de 16 ó más años)		
n_viudo	Nº de viudos en el Hogar (de 16 ó más años)	VARIABLE	DESCRIPCIÓN
n_divorciado	Nº de divorciados en el Hogar (de 16 ó más años)	cod_i	Nº de identificación del Miembro "i" del Hogar
n_espanola	Nº de españoles en el Hogar (de 16 ó más años)	sexo_i	Sexo del Miembro "i" del Hogar
n_exUE24	Nº de extranjeros del resto de la UE-25 en el Hogar (de 16 ó más años)	edad_i	Edad del Miembro "i" del Hogar
n_exResE	Nº de extranjeros del resto de Europa en el Hogar (de 16 ó más años)	gedad_i	Grupo de Edad del Miembro "i" del Hogar
n_exResM	Nº de extranjeros del resto del mundo en el Hogar (de 16 ó más años)	ned_i	Nivel de Estudios del Miembro "i" del Hogar
n_sinestud	Nº de miembros sin estudios del Hogar (de 16 ó más años)	naci_i	Nacionalidad del Miembro "i" del Hogar
n_eduprimaria	Nº de miembros con Educación Primaria del Hogar (de 16 ó más años)	cse_i	Categoría socio-económica del Miembro "i" del Hogar
n_edusec1	Nº de miembros con Educación Secundaria de 1ª etapa del Hogar (de 16 ó más años)	estcivil_i	Estado Civil del Miembro "i" del Hogar
n_edusec2	Nº de miembros con Educación Secundaria de 2ª etapa del Hogar (de 16 ó más años)	activestb1_i	Actividad laboral del Miembro "i" del Hogar (clasif.1)
n_edusuperior	Nº de miembros con Educación Superior del Hogar (de 16 ó más años)	activestb1_i_m	Actividad del establecimiento del que depende o dependía laboralmente el Miembro "i" del Hogar (clasif.1)
n_ocupado	Nº de ocupados del Hogar (de 16 ó más años)	activestb2_i	Actividad laboral del Miembro "i" del Hogar (clasif.2)
n_desempleado	Nº de activos desempleados del Hogar (de 16 ó más años)	activestb2_i_m	Actividad del establecimiento del que depende o dependía laboralmente el Miembro "i" del Hogar (clasif.2)
n_estudiante	Nº de inactivos estudiantes del Hogar (de 16 ó más años)	situprof_i	Situación profesional del Miembro "i" del Hogar
n_retirado	Nº de inactivos jubilados o retirados del Hogar (de 16 ó más años)	tipcont_i	Tipo de contrato del Miembro "i" del Hogar
n_labhogar	Nº de inactivos dedicados a labores del Hogar (de 16 ó más años)	situac3_i	Situación de Actividad del Miembro "i" del Hogar (clasif.3)
n_otrinactivo	Nº de otros inactivos del Hogar (de 16 ó más años)	ocupac_i	Ocupación del Miembro "i" del Hogar

(*) El análisis de homogeneidad de variables se ha realizado de forma independiente entre la BDH 2009 base 2013 y la EPF 2008, por una parte, y entre la BDH 2012 base 2013 y la EPF 2011, por otra, hallando resultados análogos.
Fuente: Elaboración propia a partir de las ECV base 2013 y las EPF.

Cuadro 3.5 Variables homogéneas en la BDH base 2013-EPF y en la EFF ()**

VARIABLES HOMOGÉNEAS RELATIVAS AL HOGAR			
VARIABLE	DESCRIPCIÓN	VARIABLE	DESCRIPCIÓN
fuelle	Encuesta Fuente	RNMCP_B	Renta No Monetaria Bruta por Cuenta Propia
pais	País	IIC10_RNMCP_B	Intervalo Interdecílico de la Renta No Monetaria Bruta por Cuenta Propia
ano	Año de referencia de los ingresos	PI_RNMCP_B	Percentil de la Renta No Monetaria Bruta por Cuenta Propia
anoe	Año de referencia de la encuesta	IMPRO_B	Rentas Brutas de las propiedades o terrenos
idh	Identificador del Hogar	IIC10_IMPRO_B	Intervalo Interdecílico de las Rentas Brutas de las propiedades o terrenos
w	Factor de elevación del Hogar	PI_IMPRO_B	Percentil de las Rentas Brutas de las propiedades o terrenos
wpc	Factor de elevación personal para rentas per capita	RMIX_B	Rentas Mixtas Brutas del Hogar
nmh	Nº de miembros del Hogar	IIC10_RMIX_B	Intervalo Interdecílico de las Rentas Mixtas Brutas del Hogar
nmh16	Nº de miembros del Hogar con 16 años o más	PI_RMIX_B	Percentil de las Rentas Mixtas Brutas del Hogar
nmh14	Nº de miembros del Hogar con 14 años o más	RMCA_B	Renta Bruta monetaria o cuasimonetaria de asalariados del Hogar
um	Unidad de medida de los ingresos	IIC10_RMCA_B	Intervalo Interdecílico de la Renta Bruta monetaria o cuasimonetaria de asalariados del Hogar
REGTENEVIV3_m	Régimen de tenencia de la vivienda	PI_RMCA_B	Percentil de la Renta Bruta monetaria o cuasimonetaria de asalariados del Hogar
TIPHOGAR_m	Tipo de Hogar	RNMCA_B	Renta Bruta no monetaria de asalariados del Hogar
n_dep	Nº de dependientes del Hogar (según la ECV)	IIC10_RNMCA_B	Intervalo Interdecílico de la Renta Bruta no monetaria de asalariados del Hogar
n_percep	Nº de perceptores del Hogar	PI_RNMCA_B	Percentil de la Renta Bruta no monetaria de asalariados del Hogar
n_hombre	Nº de hombres en el Hogar	RNM2CA_B	Renta Bruta no monetaria de asalariados del Hogar (coche de empresa)
n_mujer	Nº de mujeres en el Hogar	CSEFEMP	Cotizaciones sociales efectivas a cargo del empleador
n_soltero	Nº de solteros en el Hogar (de 16 ó más años)	RASAL_B	Rentas Salariales Brutas del Hogar
n_casado	Nº de casados en el Hogar (de 16 ó más años)	RASAL_B_m	Rentas Salariales Brutas del Hogar para matching
n_separado	Nº de separados en el Hogar (de 16 ó más años)	IIC10_RASAL_B_m	Intervalo Interdecílico de las Rentas Salariales Brutas del Hogar para matching
n_viudo	Nº de viudos en el Hogar (de 16 ó más años)	PI_RASAL_B_m	Percentil de las Rentas Salariales Brutas del Hogar para matching
n_divorciado	Nº de divorciados en el Hogar (de 16 ó más años)	IMCAP_B	Rentas Brutas del Capital Monetario del Hogar
n_sinestud	Nº de miembros sin estudios del Hogar (de 16 ó más años)	IIC10_IMCAP_B	Intervalo Interdecílico de las Rentas Brutas del Capital Monetario del Hogar
n_eduprimaria	Nº de miembros con Educación Primaria del Hogar (de 16 ó más años)	PI_IMCAP_B	Percentil de las Rentas Brutas del Capital Monetario del Hogar
n_edusec1	Nº de miembros con Educación Secundaria de 1º etapa del Hogar (de 16 ó más años)	RPRO_B	Rentas Brutas de la propiedad (Ingresos-Gastos Deducibles) del Hogar
n_edusec2	Nº de miembros con Educación Secundaria de 2º etapa del Hogar (de 16 ó más años)	IIC10_RPRO_B	Intervalo Interdecílico de las Rentas Brutas de la propiedad (Ingresos-Gastos Deducibles) del Hogar
n_edusuperior	Nº de miembros con Educación Superior del Hogar (de 16 ó más años)	PI_RPRO_B	Percentil de las Rentas Brutas de la propiedad (Ingresos-Gastos Deducibles) del Hogar
n_ocupado	Nº de ocupados del Hogar (de 16 ó más años)	RPRIH_B	Saldo de rentas primarias Brutas del Hogar (sin IAVP)
n_desempleado	Nº de activos desempleados del Hogar (de 16 ó más años)	RPRIH_B_m	Saldo de rentas primarias Brutas del Hogar (sin IAVP) Brutas para matching
n_estudiante	Nº de inactivos estudiantes del Hogar (de 16 ó más años)	IIC10_RPRIH_B_m	Intervalo Interdecílico del Saldo de rentas primarias Brutas del Hogar (sin IAVP) Brutas para matching
n_retirado	Nº de inactivos jubilados o retirados del Hogar (de 16 ó más años)	PI_RPRIH_B_m	Percentil del Saldo de rentas primarias Brutas del Hogar (sin IAVP) Brutas para matching
n_labhogar	Nº de inactivos dedicados a labores del Hogar (de 16 ó más años)	PSDES_B	Prestaciones por desempleo Brutas del Hogar
n_otrinactivo	Nº de otros inactivos del Hogar (de 16 ó más años)	IIC10_PSDES_B	Intervalo Interdecílico de las Prestaciones por desempleo Brutas del Hogar
RMCP_B	Renta Monetaria Bruta por Cuenta Propia	PI_PSDES_B	Percentil de las Prestaciones por desempleo Brutas del Hogar
IIC10_RMCP_B	Intervalo Interdecílico de la Renta Monetaria Bruta por Cuenta Propia	PSJUB_B	Prestaciones por jubilación / vejez Brutas del Hogar
PI_RMCP_B	Percentil de la Renta Monetaria Bruta por Cuenta Propia	IIC10_PSJUB_B	Intervalo Interdecílico de las Prestaciones por jubilación / vejez Brutas del Hogar

(**) El análisis de homogeneidad de variables se ha realizado de forma independiente entre el conjunto de datos resultante de la fusión BDH 2009 base 2013 - EPF 2008 y entre la EFF 2008, por una parte, y entre el fichero de datos producto de la fusión BDH 2012 base 2013 - EPF 2011 y entre la EFF 2011, por otra, hallando resultados análogos.

Fuente: Elaboración propia a partir de las ECV base 2013 y las EFF.

Cuadro 3.5 (cont.) Variables homogéneas en la BDH base 2013-EPF y en la EFF (**)

VARIABLES HOMOGÉNEAS RELATIVAS AL HOGAR			
VARIABLE	DESCRIPCIÓN	VARIABLE	DESCRIPCIÓN
PI_PSJUB_B	Percentil de las Prestaciones por jubilación / vejez Brutas del Hogar	PI_TCPER_B_m	Percentil de las Transferencias corrientes percibidas Brutas por el Hogar para matching
PSSUP_B	Prestaciones por supervivencia Brutas del Hogar	RT_B	Renta Total Bruta del Hogar (sin IAVP)
IIC10_PSSUP_B	Intervalo Interdecílico de las Prestaciones por supervivencia Brutas del	IIC10_RT_B	Intervalo Interdecílico de la Renta Total Bruta del Hogar (sin IAVP)
PI_PSSUP_B	Percentil de las Prestaciones por supervivencia Brutas del Hogar	PI_RT_B	Percentil de la Renta Total Bruta del Hogar (sin IAVP)
PSJYS_B_m	Prestaciones por jubilación / vejez y supervivencia Brutas para matching	I_CPRO	Indicador sobre si el Hogar tiene Ingresos por Cuenta Propia
IIC10_PSJYS_B_m	Intervalo Interdecílico de las Prestaciones por jubilación / vejez y	I_CAJE	Indicador sobre si el Hogar tiene Ingresos por Cuenta Ajena
PI_PSJYS_B_m	Percentil de las Prestaciones por jubilación / vejez y supervivencia Brutas para matching	I_PENS	Indicador sobre si el Hogar tiene Ingresos por Pensiones Contributivas y No Contributivas
PSENF_B	Prestaciones por enfermedad Brutas del Hogar	I_DES	Indicador sobre si el Hogar tiene Ingresos por Desempleo
IIC10_PSENF_B	Intervalo Interdecílico de las Prestaciones por enfermedad Brutas del Hogar	I_OSPS	Indicador sobre si el Hogar tiene Ingresos por Otras Subvenciones y Prestaciones Sociales
PI_PSENF_B	Percentil de las Prestaciones por enfermedad Brutas del Hogar	I_RCYP	Indicador sobre si el Hogar tiene Ingresos por Rentas del Capital y la
PSINV_B	Prestaciones por invalidez Brutas del Hogar	I_OTRI	Indicador sobre si el Hogar tiene Ingresos por Otros Ingresos
IIC10_PSINV_B	Intervalo Interdecílico de las Prestaciones por invalidez Brutas del Hogar	FP_ING_m	Fuente Principal de los ingresos regulares en el hogar (clasif.1)
PI_PSINV_B	Percentil de las Prestaciones por invalidez Brutas del Hogar	ORNC_B	Otras Rentas no clasificadas Brutas del Hogar
PSEEL_B	Prestaciones por enfermedad e invalidez Brutas del Hogar	GASTOH	Gasto Total del Hogar (Monetario y No Monetario)
PSEEL_B_m	Prestaciones por enfermedad e invalidez Brutas para matching	IIC10_GASTOH	Intervalo Interdecílico del Gasto Total del Hogar (Monetario y No Monetario)
IIC10_PSEEL_B_m	Intervalo Interdecílico de las Prestaciones por enfermedad e invalidez Brutas para matching	PI_GASTOH	Percentil del Gasto Total del Hogar (Monetario y No Monetario)
PI_PSEEL_B_m	Percentil de las Prestaciones por enfermedad e invalidez Brutas para matching	GCF_01	Gasto en Consumo Final del Hogar en Alimentación y Bebidas
PSFYV_B_m	Ayudas por familia/hijos, incluida Ayudas para vivienda, Brutas para matching	IIC10_GCF_01	Intervalo Interdecílico del Gasto en Consumo Final del Hogar en Alimentación y Bebidas
IIC10_PSFYV_B_m	Intervalo Interdecílico de las Ayudas por familia/hijos, incluida Ayudas para vivienda, Brutas para matching	PI_GCF_01	Percentil del Gasto en Consumo Final del Hogar en Alimentación y Bebidas
PI_PSFYV_B_m	Percentil de las Ayudas por familia/hijos, incluida Ayudas para vivienda Brutas para matching	VARIABLES HOMOGÉNEAS A LOS MIEMBROS DEL HOGAR (i=SP -Sustentador Principal- ; j=2A -Segundo Adulto en importancia económica-)	
RPEPP_B	Rentas Brutas percibidas de esquemas privados de pensiones (no SEEPROS)		
IIC10_RPEPP_B	Intervalo Interdecílico de las Rentas Brutas percibidas de esquemas privados de pensiones (no SEEPROS)		
PI_RPEPP_B	Percentil de las Rentas Brutas percibidas de esquemas privados de pensiones (no SEEPROS)	VARIABLE	DESCRIPCIÓN
PSEST_B	Ayudas para estudios Brutas del Hogar	cod_i	Número de identificación del Miembro "i"
IIC10_PSEST_B	Intervalo Interdecílico de las Ayudas para estudios Brutas del Hogar	sexo_i	Sexo del Miembro "i" del Hogar
PI_PSEST_B	Percentil de las Ayudas para estudios Brutas del Hogar	edad_i	Edad del Miembro "i" del Hogar
PSONC_B	Otras ayudas sociales al Hogar no clasificadas en otro apartado (Brutas)	gedad_i	Grupo de Edad del Miembro "i" del Hogar
PRESOC_B_m	Prestaciones Sociales Brutas al Hogar para matching	ned_i	Nivel de Estudios del Miembro "i" del Hogar
IIC10_PRESOC_B_m	Intervalo Interdecílico de las Prestaciones Sociales Brutas al Hogar para matching	cse_i	Categoría socio-económica del Miembro "i" del Hogar
PI_PRESOC_B_m	Percentil de las Prestaciones Sociales Brutas al Hogar para matching	estcivil_i	Estado Civil del Miembro "i" del Hogar
TPRDOH_B	Transferencias periódicas monetarias Brutas percibidas de otros Hogares	activestb1_i	Actividad laboral del Miembro "i" del Hogar
TCDOH_B	Transferencias Corrientes Percibidas de otros Hogares Brutas	activestb1_i_m	Actividad del establecimiento del que depende o dependía laboralmente (clasif.1) el Miembro "i" del Hogar
IIC10_TPRDOH_B	Intervalo Interdecílico de las Transferencias periódicas monetarias Brutas percibidas de otros Hogares	situprof_i	Situación profesional del Miembro "i" del Hogar
PI_TPRDOH_B	Percentil de las Transferencias periódicas monetarias Brutas percibidas de otros Hogares	tipcont_i	Tipo de contrato del Miembro "i" del Hogar
TCPER_B_m	Transferencias corrientes percibidas Brutas por el Hogar para matching	situac2_i	Situación de Actividad del Miembro "i" del Hogar (clasif.3)
IIC10_TCPER_B_m	Intervalo Interdecílico de las Transferencias corrientes percibidas Brutas por el Hogar para matching	ocupac_i	Ocupación del Miembro "i" del Hogar

(**) El análisis de homogeneidad de variables se ha realizado de forma independiente entre el conjunto de datos resultante de la fusión BDH 2009 base 2013 - EPF 2008 y entre la EFF 2008, por una parte, y entre el fichero de datos producto de la fusión BDH 2012 base 2013 - EPF 2011 y entre la EFF 2011, por otra, hallando resultados análogos.

Fuente: Elaboración propia a partir de las ECV base 2013 y las EFF.

3.1.2.3. Selección de variables de fusión

Una vez que hemos podido disponer de variables homogéneas en sendas encuestas, el paso siguiente que tenemos que realizar consiste en seleccionar las variables que finalmente intervendrán en el proceso de *matching*, es decir, aquellas que nos van a permitir emparejar registros en las muestras que pretendemos fusionar. Ésta es una decisión importante en el proceso de fusión estadística, puesto que es evidente que la calidad de la información de la muestra resultante depende, en buena medida, de las variables que intervienen a la hora de enlazar las observaciones. Debemos tener en cuenta que la elección de una variable de fusión determinada estará condicionada por el tipo de análisis que deseemos efectuar con la muestra fusionada. En particular, hay que tener en cuenta qué variables, de entre todas las que integran las muestras objeto de fusión (Y, X, Z), declaramos objetivo principal de nuestro estudio. A partir de aquí, no existe un método claro que proporcione una guía para seleccionar estas variables, y a veces las decisiones son subjetivas, e incluso, basadas en criterios *ad hoc* relacionados con las posibilidades de la información con la que trabajamos⁵⁰.

Las variables finalmente seleccionadas para efectuar el emparejamiento de las observaciones, de entre el conjunto de variables comunes descritas en los cuadros 3.4 y 3.5, deben cumplir dos criterios⁵¹. En primer lugar, debe haber similitud (coherencia) en las distribuciones de dichas variables en los conjuntos de datos a fusionar. En segundo lugar, las variables deben ser significativas para explicar las variaciones en las variables objetivos a incorporar mediante la fusión.

Coherencia de las distribuciones

La literatura pone de relieve dos métodos principales para calcular el grado en que las distribuciones de las variables son similares en los conjuntos de datos a fusionar. La primera consiste en realizar una comparación simple de las distribuciones marginales de frecuencias para cada una de las variables potenciales de fusión entre las distribuciones de probabilidad del conjunto de datos del donante y del receptor. La segunda se basa en el uso de una medida de distancia entre ellas, denominada Distancia de Hellinger (HD).

La medida HD proporciona un solo valor como medida de la similitud entre las dos distribuciones de probabilidad de una misma variable en los distintos conjuntos de datos a fusionar. Así, dadas dos distribuciones discretas de probabilidad, $V=(v_1, v_2, \dots, v_D)$ y $V'=(v'_1, v'_2, \dots, v'_R)$, la Distancia de Hellinger se define como:

⁵⁰ En muchas ocasiones, la elección de las variables de fusión puede estar condicionada por razones ligadas a errores de medida y/o representatividad. Por ejemplo, en una fusión entre datos fiscales y las familias representadas en las encuestas de presupuestos familiares (Calonge y Manresa, 1997), el nivel de salarios y la categoría de asalariados constituyen variables claves; no sólo porque la masa salarial es el componente más importante de la renta, sino también porque es la variable de renta que presenta una mejor fiabilidad en las encuestas de presupuestos.

⁵¹ Inicialmente, hay que señalar que se han controlado el número de valores missing de cada una de las variables homogéneas candidatas a la fusión, procediendo a estudiar, en primer lugar, la coherencia de las distribuciones de aquellas variables cuyo número de valores missing sea inferior al 25%, siendo en la práctica totalidad de las variables consideradas mucho menor.

$$HD(V, V') = \sqrt{\frac{1}{2} \sum_{i=1}^K (\sqrt{p(V=i)} - \sqrt{p(V'=i)})^2} = \sqrt{\frac{1}{2} \sum_{i=1}^K \left(\sqrt{\frac{n_{Di}}{N_D}} - \sqrt{\frac{n_{Ri}}{N_R}} \right)^2}$$

donde,

K es el número total de categorías distintas que puede presentar la variable;

n_{Di} es la frecuencia de la categoría i en el fichero de datos donante (D);

n_{Ri} es la frecuencia de la categoría i en el fichero de datos receptor (R);

N_D es el total de observaciones del conjunto de datos D;

N_R es el total de observaciones de R.

Un valor HD (%) del 0% indica una perfecta similitud entre las dos distribuciones probabilísticas, mientras que un valor HD del 100% indica una discrepancia máxima.

Desafortunadamente, no es posible establecer objetivamente un umbral de valores aceptables de la distancia, según el cual las dos distribuciones se pueden decir próximas. De forma recurrente, la literatura considera dos distribuciones cercanas si la Distancia de Hellinger es inferior al 5%; en caso contrario, debe plantear preocupaciones sobre las similitudes en las distribuciones. Sin embargo, esta es simplemente una regla general sin muchos antecedentes teóricos, y el umbral establecido es, por otro lado, arbitrario (Leulescu *et al.*, 2011).

En nuestro estudio, la similitud de las variables comunes se ha analizado a través del cálculo de las Distancias de Hellinger, para las diferentes modalidades que presenta la variable *Tipo de hogar para matching* (TIPHOGAR_m), entre las distribuciones de la BDH base 2013 y las de la EPF, por un lado, y entre las mismas distribuciones de la BDH base 2013 y las distribuciones de la EFF, por otro. A modo de ejemplo, los cuadros 3.6 y 3.7 muestran los valores HD obtenidos para los *hogares de dos adultos sin niños dependientes económicamente, teniendo ambos menos de 65 años* (TIPHOGAR_m=8), entre las distribuciones de la BDH base 2013 año 2009 y las de la EPF 2008, y las de la primera respecto de las de la EFF 2008, y entre las de la BDH base 2013 año 2012 y las de la EPF 2011, y las de la BDH base 2013 año 2012 respecto de las de la EFF 2011, respectivamente. Considerando el 7% como referencia para la cercanía de dos distribuciones, en general, las distribuciones de variables demográficas básicas referidas al número de miembros del hogar están bastante bien alineadas tanto en BDH base 2013 y en EPF como en BDH base 2013 y en EFF, para los años 2008 y 2011, mientras que se observan discrepancias más altas, con HD superiores al 7%, en el régimen de tenencia de la vivienda principal del hogar, en las variables relacionadas con la información laboral referidas tanto al sustentador principal como al segundo adulto, y en algunos indicadores de ciertos tipos de ingresos.

Análisis del poder explicativo de las variables comunes

Para todas aquellas variables cuyos valores HD se encuentran por debajo del umbral considerado, el segundo criterio aplicado para la selección de las variables comunes es que deben mostrar una buena asociación con las

variables objetivos a fusionar (variables no recogidas en el archivo receptor, sólo en el donante, y que se necesita incorporar en el archivo receptor). La situación óptima, cuando las variables comunes contienen toda la información compartida por dichas variables objetivo y por lo tanto cumplen con la suposición de independencia condicional (CIA).

Por lo tanto, con el fin de hacer una selección adicional de variables comunes que se utilizarán en el proceso de fusión, su asociación con las variables objetivo de los conjuntos de datos a fusionar (el *gasto total del hogar*, para la fusión entre la BDH base 2013 y la EPF, y la *riqueza neta del hogar* para la fusión entre la BDH base 2013-EPF y la EFF) se analizó utilizando el Coeficiente V de Cramer (D 'Orazio *et al.*, 2006), derivado de la prueba de Chi-cuadrado de Pearson⁵², el cual permite obtener un índice con valor máximo (que indica la mayor asociación entre variables) igual a 1 (el valor mínimo es 0, que indica no asociación). En los cuadros 3.6 y 3.7 se presentan los resultados del Coeficiente V de Cramer para las variables comunes cuyas distribuciones pueden considerarse similares en base a los valores HD comentados anteriormente y que muestran una asociación significativa con las variables objetivo al nivel del 1% (p-valores⁵³<0,01) para los *hogares de dos adultos sin niños dependientes económicamente, teniendo ambos menos de 65 años*, a modo de ejemplo (estos cuadros han sido calculadas y consideradas para todos los tipos de hogares). En concreto, en dichas cuadros puede observarse cómo las variables que tienen dato en la columna "V de Cramer" son las variables que se utilizarán para la fusión de este tipo de hogares y se caracterizarán por tener un HD inferior al 7% (en caso de no cumplir esta condición se identifican con fondo gris en la columna "Hd(%)") y, además, por presentar una asociación significativa con la variable objetivo al 1% (en caso contrario, los datos de la columna "sig." que muestra dicho nivel de significación también están identificadas con fondo gris). Los resultados muestran que, en general, dichas variables comunes tienden a estar asociadas con las variables objetivo y parecen ser buenos candidatos como variables de control para el emparejamiento de registros.

⁵² Para medir la asociación entre dos variables categóricas, X e Y, existen varias medidas basadas en el clásico estadístico Chi-cuadrado de Pearson:

$$X^2 = \sum_{i=1}^I \sum_{j=1}^J \frac{(n_{ij} - m_{ij})^2}{m_{ij}}$$

donde n_{ij} y m_{ij} representan las frecuencias observadas y esperadas, respectivamente, $i = 1, \dots, I$, $j = 1, \dots, J$. Una de las medidas de asociación más útil es la V de Cramer:

$$V = \sqrt{\frac{\frac{X^2}{n}}{\min\{I-1, J-1\}}}$$

De hecho, esta medida tiene la importante característica de asumir valores entre 0 (ausencia de asociación entre X e Y) y 1 (asociación perfecta entre X e Y).

⁵³ Probabilidad de rechazar la hipótesis nula de ausencia de asociación, cuando es cierta, con la muestra actual u otra más desfavorable..

Cuadro 3.6 Selección de variables de *matching* en hogares de dos adultos sin niños dependientes económicamente para el año de referencia de los ingresos 2008

Fusión BDH base 2013 año 2009 vs. EPF año 2008				Fusión BDH base 2013 año 2009 -EPF año 2008 vs. EFF año 2008			
Variables comunes	HD (%)	V de Cramer (GASTOH)	sign	Variables comunes	HD (%)	V de Cramer (riquezane)	sign
nmh	0,0%			nmh	0,0%		
nmh16	0,0%			nmh16	0,0%		
nmh14	0,0%			nmh14	0,0%		
sexo_SP	8,8%		0,725	sexo_SP	4,1%	0,184	0,000
gedad_SP	4,6%	0,102	0,000	gedad_SP	8,4%	0,273	0,000
ned_SP	6,6%	0,188	0,000	ned_SP	5,9%	0,163	0,000
naci_SP	5,9%	0,203	0,000	cse_SP	13,1%		0,000
cse_SP	15,7%		0,000	estcivil_SP	9,0%		0,000
estcivil_SP	3,6%	0,094	0,000	situprof_SP	5,5%	0,161	0,000
situprof_SP	6,9%	0,111	0,000	tipcont_SP	1,7%	0,202	0,000
tipcont_SP	12,4%		0,000	situac2_SP	8,8%		0,000
situac3_SP	10,2%		0,000	ocupac_SP	10,0%		0,000
ocupac_SP	11,0%		0,000	REGTENEVIV3_m	9,4%		0,000
REGION_m	2,8%	0,104	0,000	TIPHOGAR_m	0,0%		
HAB_m	4,2%	0,098	0,000	n_dep	0,0%		
TIPOVIV_m	4,7%	0,101	0,000	n_percep	10,8%		0,000
REGTENEVIV_m	8,0%		0,000	n_hombre	4,6%	0,154	0,000
NUMHABIT_m	3,9%	0,099	0,000	n_mujer	4,6%	0,154	0,000
n_dep	0,0%			n_soltero	7,7%		0,000
n_percep	9,2%		0,000	n_casado	9,8%		0,000
n_hombre	4,9%	0,115	0,000	n_separado	2,0%	0,136	0,000
n_mujer	4,9%	0,115	0,000	n_viudo	2,7%	0,179	0,000
n_soltero	3,5%	0,118	0,000	n_divorciado	5,7%	0,151	0,000
n_casado	3,7%	0,107	0,000	n_sinestud	5,6%	0,116	0,000
n_separado	2,5%	0,085	0,001	n_eduprimaria	8,1%		0,000
n_viudo	3,0%	0,092	0,000	n_edusec1	5,3%	0,202	0,000
n_divorciado	3,1%	0,077	0,009	n_edusec2	3,0%	0,124	0,000
n_espanola	4,9%	0,198	0,000	n_edusuperior	1,9%	0,206	0,000
n_exUE24	1,0%	0,085	0,001	n_ocupado	2,5%	0,161	0,000
n_exResE	5,6%	0,219	0,000	n_desempleado	2,3%	0,223	0,000
n_exResM	0,0%			n_estudiante	2,6%	0,200	0,000
n_sinestud	7,3%		0,000	n_retirado	6,8%	0,172	0,000
n_eduprimaria	4,0%	0,173	0,000	n_labhogar	0,8%	0,246	0,000
n_edusec1	4,6%	0,132	0,000	n_otrinactivo	12,9%		0,000
n_edusec2	4,6%	0,082	0,002	I_CPRO	3,0%	0,319	0,000
n_edusuperior	1,5%	0,256	0,000	I_CAJE	9,8%		0,000
n_ocupado	5,9%	0,182	0,000	I_PENS	0,8%	0,233	0,000
n_desempleado	6,5%	0,189	0,000	I_DES	14,5%		0,000
n_estudiante	2,5%	0,104	0,000	I_OSPS	5,8%	0,167	0,000
n_retirado	3,2%	0,082	0,002	I_RCYP	9,1%		0,000
n_labhogar	0,2%	0,133	0,000	I_OTRI	5,5%	0,238	0,000
n_otrinactivo	9,9%		0,000	FP_ING_m	9,8%	0,204	0,000
IIC10_INTOTALES_m	7,0%	0,226	0,000	sexo_2A	1,6%	0,177	0,000
IIC10_lintotales_m	7,0%		0,000	gedad_2A	8,2%		0,000
IIC10_IND1_IT_m	7,0%	0,226	0,000	estcivil_2A	8,7%		0,000
I_CPRO	7,1%		0,000	ned_2A	9,1%		0,000
I_CAJE	3,7%	0,136	0,000	cse_2A	12,3%		0,000
I_PENS	6,3%	0,086	0,009	ocupac_2A	10,1%		0,000
I_DES	20,6%		0,000	situprof_2A	10,6%		0,000
I_OSPS	21,5%		0,000	tipcont_2A	1,9%	0,179	0,000
I_RCYP	41,6%		0,000	situac2_2A	9,0%		0,000
I_OTRI	4,3%		0,077	activestb1_SP_m	9,8%		0,000
FP_ING_m	15,2%		0,000	activestb1_2A_m	10,3%		0,000
FP_ing2_m	12,0%		0,000	IIC10_RMCP_B	4,8%	0,123	0,000
sexo_2A	5,4%	0,089	0,005	IIC10_RNMCP_B	4,1%	0,062	0,003
gedad_2A	5,4%	0,094	0,000	IIC10_RMIX_B	6,1%	0,125	0,000
estcivil_2A	4,5%	0,081	0,000	IIC10_RMCA_B	6,8%	0,149	0,000
ned_2A	8,0%		0,000	IIC10_RNMCA_B	5,1%	0,082	0,000
naci_2A	4,9%	0,198	0,000	IIC10_RASAL_B_m	6,4%	0,154	0,000
cse_2A	31,8%		0,000	IIC10_IMCAP_B	5,5%	0,190	0,000
ocupac_2A	21,8%		0,000	IIC10_IMPRO_B	4,1%	0,116	0,000
situprof_2A	20,3%		0,000	IIC10_RPRO_B	5,1%	0,194	0,000
tipcont_2A	25,1%		0,001	IIC10_RPRIH_B_m	9,9%		0,000
situac3_2A	8,7%		0,000	IIC10_PSDES_B	5,6%	0,096	0,000
activestb1_SP_m	18,1%		0,000	IIC10_PSJUB_B	15,4%		0,000
activestb1_2A_m	23,5%		0,000	IIC10_PSSUP_B	14,4%		0,000
				IIC10_PSJYS_B_m	7,1%		0,000
				IIC10_PSENF_B	3,4%		0,479
				IIC10_PSINV_B	3,8%	0,078	0,000
				IIC10_PSEEL_B_m	5,5%	0,076	0,000
				IIC10_PSEST_B	4,1%	0,067	0,000
				IIC10_PSFYV_B_m	3,7%	0,061	0,005
				IIC10_RPEPP_B	3,5%		0,773
				IIC10_PRESOC_B_m	12,6%	0,138	0,000
				IIC10_TPRDOH_B	3,3%	0,071	0,000
				IIC10_TCPER_B_m	14,1%		0,000
				IIC10_RT_B	9,9%	0,236	0,000
				IIC10_GASTOH	6,8%	0,193	0,000
				IIC10_GCF_01	6,9%	0,169	0,000

Fuente: Elaboración propia a partir de la ECV base 2013 año 2009 y las EPF y EFF año 2008.

Cuadro 3.7 Selección de variables de *matching* en hogares de dos adultos sin niños dependientes económicamente para el año de referencia de los ingresos 2011

Fusión BDH base 2013 año 2012 vs. EPF año 2011				Fusión BDH base 2013 año 2012 -EPF año 2011 vs. EFF año 2011			
Variables comunes	HD (%)	V de Cramer (GASTOH)	sign	Variables comunes	HD (%)	V de Cramer (riquezanel)	sign
nmh	0,0%			nmh	0,0%		
nmh16	0,0%			nmh16	0,0%		
nmh14	0,0%			nmh14	0,0%		
sexo_SP	3,6%		0,094	sexo_SP	2,7%	0,141	0,000
gedad_SP	3,1%	0,069	0,010	gedad_SP	10,1%	0,331	0,000
ned_SP	11,9%		0,000	ned_SP	11,6%		0,000
naci_SP	3,8%	0,193	0,000	cse_SP	12,0%		0,000
cse_SP	15,4%		0,000	estcivil_SP	9,9%		0,000
estcivil_SP	4,6%	0,092	0,000	situprof_SP	2,8%	0,244	0,000
situprof_SP	7,1%		0,000	tipcont_SP	0,4%	0,235	0,000
tipcont_SP	10,3%		0,000	situac2_SP	12,3%		0,000
situac3_SP	10,8%		0,000	ocupac_SP	10,7%		0,000
ocupac_SP	9,5%		0,000	REGTENEVIV3_m	7,8%		0,000
REGION_m	3,4%	0,126	0,000	TIPHOGAR_m	0,0%		
HAB_m	1,5%	0,094	0,000	n_dep	0,0%		
TIPOVIV_m	6,4%	0,093	0,000	n_percep	6,8%	0,153	0,000
REGTENEVIV_m	11,3%		0,000	n_hombre	4,3%	0,165	0,000
NUMHABIT_m	6,6%	0,121	0,000	n_mujer	4,3%	0,165	0,000
n_dep	0,0%		0,000	n_soltero	8,9%		0,000
n_percep	8,6%		0,000	n_casado	9,6%		0,000
n_hombre	1,7%		0,021	n_separado	3,4%	0,114	0,000
n_mujer	1,7%		0,021	n_viudo	0,6%	0,163	0,000
n_soltero	4,5%	0,104	0,000	n_divorciado	3,8%	0,178	0,000
n_casado	4,6%	0,109	0,000	n_sinestud	4,9%	0,151	0,000
n_separado	3,1%	0,100	0,000	n_eduprimaria	8,1%		0,000
n_viudo	3,2%	0,080	0,003	n_edusec1	5,1%	0,236	0,000
n_divorciado	2,0%		0,030	n_edusec2	5,6%	0,186	0,000
n_espanola	3,4%	0,200	0,000	n_edusuperior	2,6%	0,227	0,000
n_exUE24	2,2%	0,133	0,000	n_ocupado	5,0%	0,188	0,000
n_exResE	5,6%	0,187	0,000	n_desempleado	6,3%	0,261	0,000
n_exResM	0,0%			n_estudiante	4,5%	0,213	0,000
n_sinestud	5,9%	0,087	0,000	n_retirado	10,5%		0,000
n_eduprimaria	11,8%		0,000	n_labhogar	1,0%	0,260	0,000
n_edusec1	9,6%		0,000	n_otrinactivo	15,4%		0,000
n_edusec2	2,2%		0,023	L_CPRO	1,2%	0,295	0,000
n_edusuperior	4,1%	0,268	0,000	L_CAJE	8,8%		0,000
n_ocupado	5,7%	0,234	0,000	L_PENS	1,3%	0,301	0,000
n_desempleado	4,2%	0,248	0,000	L_DES	12,7%		0,000
n_estudiante	2,4%		0,015	L_OSPS	3,6%	0,107	0,000
n_retirado	1,3%		0,700	L_RCYP	8,8%		0,000
n_labhogar	1,6%	0,090	0,000	L_OTRI	12,0%		0,000
n_otrinactivo	10,8%		0,113	FP_ING_m	8,2%	0,219	0,000
IIC10_INTOTALES_m	4,3%	0,243	0,000	sexo_2A	5,5%	0,216	0,000
IIC10_lintotales_m	4,3%	0,243	0,000	gedad_2A	11,8%		0,000
IIC10_IND1_IT_m	4,3%	0,243	0,000	estcivil_2A	9,0%		0,000
L_CPRO	3,9%	0,085	0,010	ned_2A	8,1%		0,000
L_CAJE	6,3%	0,235	0,000	cse_2A	12,8%		0,000
L_PENS	5,4%		0,253	ocupac_2A	10,3%		0,000
L_DES	17,1%		0,000	situprof_2A	9,3%		0,000
L_OSPS	18,6%		0,030	tipcont_2A	2,9%	0,182	0,000
L_RCYP	37,7%		0,000	situac2_2A	7,4%		0,000
L_OTRI	3,4%		0,139	IIC10_RMCP_B	5,8%	0,117	0,000
FP_ING_m	14,7%		0,000	IIC10_RNMCP_B	5,4%		0,032
FP_ing2_m	9,7%	0,161	0,000	IIC10_RMIX_B	4,8%	0,119	0,000
sexo_2A	3,1%		0,020	IIC10_RMCA_B	6,7%	0,182	0,000
gedad_2A	4,0%	0,082	0,000	IIC10_RNMCA_B	5,3%	0,096	0,000
estcivil_2A	3,6%	0,090	0,000	IIC10_RASAL_B_m	6,6%	0,181	0,000
ned_2A	12,7%		0,000	IIC10_IMCAP_B	4,3%	0,205	0,000
naci_2A	6,1%	0,179	0,000	IIC10_IMPRO_B	5,0%	0,084	0,000
cse_2A	35,5%		0,000	IIC10_RPRO_B	4,3%	0,190	0,000
ocupac_2A	23,3%		0,000	IIC10_RPRIH_B_m	7,5%	0,211	0,000
situprof_2A	21,4%		0,002	IIC10_PSDS_B	5,4%	0,130	0,000
tipcont_2A	25,6%		0,000	IIC10_PSJUB_B	7,0%	0,131	0,000
situac3_2A	8,9%		0,000	IIC10_PSSUP_B	11,1%		0,000
activestb2_SP_m	10,2%		0,000	IIC10_PSJYS_B_m	9,7%		0,000
activestb1_SP_m	6,7%	0,131	0,000	IIC10_PSENF_B	5,4%		0,221
activestb2_2A_m	17,8%		0,000	IIC10_PSINV_B	4,6%	0,068	0,001
activestb1_2A_m	16,0%		0,000	IIC10_PSEEL_B_m	4,2%	0,069	0,000
				IIC10_PSEST_B	5,2%		0,018
				IIC10_PSFYV_B_m	5,4%		0,381
				IIC10_RPEPP_B	5,6%		0,466
				IIC10_PRESOC_B_m	9,1%	0,170	0,000
				IIC10_TPRDOH_B	5,5%	0,078	0,000
				IIC10_TCPER_B_m	9,5%		0,000
				IIC10_RT_B	5,5%	0,252	0,000
				IIC10_GASTOH	7,3%	0,228	0,000
				IIC10_GCF_01	4,6%	0,213	0,000

Fuente: Elaboración propia a partir de la ECV base 2013 año 2012 y las EPF y EFF año 2012.

3.1.2.4. Elección del tipo de fusión y del criterio de función de distancia a aplicar

Una vez tenemos las muestras convenientemente homogeneizadas y hemos seleccionado las variables de fusión, el siguiente paso consiste en “emparejar” cada observación de la muestra base con una observación de la muestra no-base. Dado un elemento de la muestra base, el resultado de la fusión vendrá dado al seleccionar como par aquel que presente la menor distancia, de entre todos los emparejamientos potenciales.

En la literatura existen dos tipos de procedimientos de fusión, que se denominan fusión *no-restringida* y fusión *restringida*, dependiendo de si existe o no reemplazamiento de las observaciones.

La fusión *no-restringida* se caracteriza por fijar un conjunto de datos como muestra base o muestra receptora. A cada observación de esta muestra le asignamos una observación de la muestra no-base o muestra donante, de tal forma que algunos registros de esta última muestra formarán parte de la fusión resultante y otros no. Incluso, es posible la repetición de las observaciones no-base en el proceso de asignación. Este tipo de fusión utiliza el criterio de mínima distancia garantizando así el “emparejamiento” de los registros más próximos. Sin embargo, la distribución final no conserva las mismas propiedades estadísticas de la muestra no-base previa a la fusión. Ahora bien, esto no tiene por qué ser un objetivo en sí mismo. Aunque no se puede obtener en la muestra fusionada o muestra *match* las distribuciones exactas para las variables (X, Z) , el procedimiento no-restringido será adecuado, desde una perspectiva de inferencia, si no existen diferencias significativas entre las distribuciones de las variables de la muestra resultante y la muestra de control.

El método *restringido* de fusión estadística tiene la propiedad interesante de preservar las distribuciones conjuntas y marginales de las variables (X, Z) de la muestra no-base, puesto que todas las observaciones de esta muestra estarán presentes en el fichero final resultante.

Una fusión restringida consiste en minimizar una función de distancia sujeta a un conjunto de restricciones. Si representamos por d_{ij} la distancia entre los elementos i, j en los ficheros base y no-base, respectivamente, y por w_{ij} los pesos asignados a cada par de elementos $[i, j]$ en la muestra resultante de la fusión, la presentación más estricta de fusión *restringida* se formaliza como un problema de minimización con restricciones:

$$\text{Min} \sum_{i=1}^{n_{\text{base}}} \sum_{j=1}^{n_{\text{no-base}}} w_{ij} d_{ij} [X_i^{\text{base}} X_j^{\text{no-base}}]$$

sujeto a:

$$\begin{aligned} \sum_{i=1}^{n_{\text{no-base}}} w_{ij} &= W_i \quad ; \quad i = 1, \dots, n_{\text{base}} \quad W_i > 0 \\ \sum_{j=1}^{n_{\text{base}}} w_{ij} &= W_j \quad ; \quad j = 1, \dots, n_{\text{no-base}} \quad W_j > 0 \end{aligned}$$

Si el número de restricciones es relativamente alto (igual a la suma del número de observaciones entre ambas muestras), este procedimiento puede tener un coste elevado en cuanto a computación. Sin embargo, en la literatura encontramos otros métodos de fusión *restringida* que, manteniendo sus propiedades (la misma distribución de las variables antes y después de la fusión), no constituyen, de forma explícita, un problema de optimización con restricciones. Básicamente, el procedimiento consiste en ordenar a los individuos en función del rango de una o más variables continuas. Dado que no será usual que los pesos o factores de elevación coincidan entre ambos conjuntos de datos, será necesario replicar las observaciones y *expandir* las muestras antes de realizar la fusión de las observaciones.

La expansión tiene como objetivo enlazar observaciones de tal forma que tengan la misma ponderación antes de ser fusionadas.

El resultado final de esta expansión es que la muestra fusionada incrementará en número de observaciones con relación a la muestra base. Ahora bien, a diferencia del tipo de fusión anterior, por ejemplo, la media ponderada de una determinada variable en el fichero final será la misma que la de los datos previos a la fusión. Sin embargo, este tipo de procedimiento, al contrario de los que ocurre con la fusión *no-restringida*, no asegura una asignación basada en el criterio de menor distancia.

La fusión *no-restringida* pone énfasis en la distancia entre dos observaciones, y la representatividad de las variables Z tiene un papel secundario. Por el contrario, la fusión *restringida* respeta la distribución de estas variables, sin garantizar el criterio de mínima distancia entre observaciones. Esto último puede suponer, en la práctica, penalizar la estimación vía regresión de relaciones del tipo $Y-X_{\text{NO-BASE}}$ o $Y-Z$, por ejemplo.

La elección del tipo de fusión es una decisión de índole práctica. La fusión *restringida* es un procedimiento de cierta complejidad de aplicación cuando tenemos un gran número de observaciones y varias variables de fusión. La fusión *no-restringida* tiene la ventaja de una mayor facilidad práctica y puede ser, en general, un método de fusión apropiado siempre y cuando se contraste, de alguna manera, la bondad de la fusión efectuada. Sin embargo, en las regiones o partes del espacio muestral con poca densidad, donde el número de observaciones de la muestra base es pequeño, la fusión *no-restringida* será probablemente poco representativa de la distribución de las variables Z .

Para llevar a cabo el proceso de fusión de la BDH base 2013, la ECV y la EPF emplearemos un método mixto, como se muestra en el programa que lo realiza presentado en el Anexo 2.

El proceso comienza reponderando proporcionalmente los pesos de la fuente donante para que la suma de pesos en ambas fuentes coincida exactamente. Y, como en los métodos no restringidos, el proceso se inicia buscando el caso donante más próximo al primer caso base considerado. Si la ponderación del caso no base es superior a la del caso base, el caso base satura toda su ponderación con el caso donante el caso base y se fusiona con el donante; pero el caso donante generaría un nuevo caso donante con las mismas características y la parte de su ponderación no empleada como nueva ponderación, estaría disponible para usar en otras fusiones. Si, por el

contrario, la ponderación del caso base es superior a la del caso donante, el caso base se expandiría en dos, como en los métodos restringidos: uno ya fusionado con el caso donante pero con la ponderación que heredaría de él, y otro disponible para ser nuevamente fusionado, con una ponderación igual al resto de ponderación aun no saturada (ponderación inicial menos ponderación del donante).

Este proceso genera un mayor número de casos que los inicialmente existentes en el archivo maestro, y como ocurre en todos los procedimientos restringidos, en este proceso mixto también ocurre que a medida que avanzamos en el proceso de fusiones, los casos óptimos para fusionar van siendo cada vez menos parecidos. Así que, para evitar estos dos inconvenientes (el primero de incremento quizás poco eficiente de la complejidad computacional y el segundo de pérdida forzada de representatividad de los casos), el proceso mixto que proponemos presenta dos parámetros que permiten flexibilizar su uso. El parámetro "umbral" permite reiterar el proceso de búsqueda de nuevos casos siempre que su disimilaridad sea inferior a una proporción aceptada de la máxima teóricamente observable (en nuestro caso lo hemos fijado en el 20% de la máxima). Y el parámetro "nciclos" permite limitar la búsqueda de nuevos casos para evitar el crecimiento desmesurado de la nueva base fusionada (en nuestro caso, hemos limitado el número máximo de expansiones de un caso a 9). Si el caso disponible más próximo estuviera a mayor distancia de la permitida o el número de expansiones de un caso hubiera llegado a su límite, la ponderación restante se repartiría proporcionalmente entre las expansiones que haya tenido el caso base, como en los métodos no restringidos, quedando así totalmente saturado y fusionado.

Una vez decidido el tipo de fusión a realizar, tenemos que adoptar un criterio de distancia que determina la asignación entre las observaciones de los distintos conjuntos de datos. Existen muchas formas de realizar el enlace entre dos observaciones teniendo en cuenta que las variables de fusión pueden ser cualitativas y/o numéricas. Sin embargo, todos los métodos empleados se pueden reducir básicamente a dos. El primero de ellos consiste en crear tipologías de individuos *similares* mediante el cruzamiento de las categorías de las variables cualitativas. En una segunda etapa, la asignación entre observaciones se realiza dentro de cada grupo mediante el cálculo, por ejemplo, de la distancia euclídea (u otro tipo de distancia) definida sobre una o más variables cuantitativas (el salario, la edad, etc.). A partir de aquí, se elige como fusión de una observación aquel par con la menor distancia entre todos los emparejamientos potenciales.

Otra opción consiste en definir, desde un principio, una función de distancia apropiada sobre todo el vector de variables categóricas y numéricas. Éste es el método que proponemos en el presente trabajo. Así pues, como medida de proximidad, para evaluar la afinidad de los hogares encuestados en la ECV y en la EPF, y en la base de hogares resultante y en la EFF, hemos propuesto y empleado la siguiente medida de disimilitud, basada en el Coeficiente de similaridad de Gower⁵⁴, que es una función de distancia que se puede aplicar sobre todo tipo de variables, sean éstas cuantitativas o no. Esta distancia se define por:

⁵⁴ Gower (1971).

$$d_{ij} = \frac{\sum_{k=1}^P W_k \delta_{ijk} D_{ijk}}{\sum_{k=1}^P W_k \delta_{ijk}}$$

donde,

D_{ijk} representa el coeficiente de similaridad entre el individuo i –ésimo de una muestra y el individuo j –ésimo de la otra muestra, tal que:

$$D_{ijk} = \begin{cases} \left| \frac{x_{ik} - x_{jk}}{\max_h \{x_{hk}\} - \min_h \{x_{hk}\}} \right|, & \text{si } X_k \text{ es métrica} \\ \begin{cases} 0, & \text{si } x_{ik} = x_{jk} \\ 1, & \text{si } x_{ik} \neq x_{jk} \end{cases}, & \text{si } X_k \text{ es atributo} \end{cases}$$

p representa el tamaño del vector $X(\cdot)$ de variables de fusión;

W_k es el factor de ponderación que se aplica a cada variable k –ésima de fusión dándoles un peso específico -unas son más importantes que otras- a la hora de enlazar ambas bases de datos; en la práctica la elección de estos pesos es generalmente subjetiva, y no existe un procedimiento óptimo para determinarlos⁵⁵;

δ_{ijk} es el coeficiente que mide la posibilidad de hacer comparaciones sobre una característica k –ésima para i y j , tal que:

$$\delta_{ijk} = \begin{cases} 1, & \text{si ambos } x_{ik} \text{ y } x_{jk} \text{ están disponibles} \\ d_k \in [0,1], & \text{si } x_{ik} \text{ está disponible y } x_{jk} \text{ no, o viceversa} \\ 0, & \text{si ni } x_{ik} \text{ ni } x_{jk} \text{ están disponibles} \end{cases}$$

En los datos utilizados en nuestra aplicación hemos considerado $\delta_{ijk} = 0,5$, es decir, que la presencia de una característica en un caso y la ausencia de la misma en el otro -por ejemplo, dos hogares donde en uno se sabe los hijos que tiene y en el otro no- se penaliza con 50% de la máxima distancia que teóricamente podría observarse en dicha variable.

En los cuadros 3.8 a 3.11 se muestran, para cada clase de hogares determinada por las diferentes modalidades que presenta la variable *Tipo de hogar para matching* (TIPHOGAR_m), las variables comunes obtenidas de acuerdo al análisis realizado en el apartado 3.1.2.3., junto a sus factores de comparación y de ponderación, las cuales participan en cada proceso de fusión que llevaremos a cabo, es decir, entre la BDH base 2013 y la EPF, y entre base de datos resultante y la EFF, para estas encuestas cuyos ingresos se recogen para los años 2008 y 2011, respectivamente.

A la izquierda de estos cuadros, se presentan las variables comunes consideradas cuyas distribuciones son coherentes en base a los valores HD comentados anteriormente y cuya V de Cramer mostró una asociación significativa de dichas variables comunes con las variables objetivo al nivel del 1% para cada tipo de hogar. Se han estructurado en bloques de filas diferenciados de acuerdo con si se refieren a características del sustentador principal, o a características del segundo adulto principal, o si se refieren de

⁵⁵ Como más adelante veremos, en este trabajo hemos optado por calcularlos a través del Coeficiente V de Cramer (multiplicado por 1000 y por un factor de ponderación que enfatiza su importancia para el análisis).

forma resumida a la composición del hogar (nombradas con el prefijo 'n' como indicativo del "número de miembros" con tal característica), o si se refieren a aspectos socio-demográficos del hogar, o si se corresponden con variables económicas del hogar (relativas a los ingresos, gastos, etc). Por otro lado, los once tipos diferentes de hogar, que determinan las once clases de hogares analizadas, se recogen en las sucesivas columnas del cuadro: *Una persona de menos de 30 años (1); Una persona de entre 30 y 64 años (2); Una persona de 65 o más años (3); 2 adultos sin niños dependientes económicamente, al menos una persona de 65 o más años (7); 2 adultos sin niños dependientes económicamente, teniendo ambos menos de 65 años (8); Otros hogares sin niños dependientes económicamente (9); Un adulto con al menos un niño dependiente (10); Dos adultos con un niño dependiente (11); Dos adultos con dos niños dependientes (12); Dos adultos con tres o más niños dependientes (13); Otros hogares con niños dependientes (14)*. Y, en el interior del cuadro, se muestran aquellas variables comunes cuyas distribuciones son coherentes según la medida HD y cuya V de Cramer mostró una asociación significativa de dichas variables comunes con las variables objetivo al nivel del 1% para cada tipo de hogar, recogiendo en el cruce de sus correspondientes fila y columna el valor absoluto del Coeficiente V de Cramer (multiplicado por 1000 y por un factor de ponderación que enfatiza su importancia para el análisis⁵⁶) que presentaron con relación a la variable *GASTOH* en el caso de la fusión BDHbase2013-EPF (cuadros 3.8 y 3.9), y a la variable *riquezane1* en la fusión BDHbase2013-EPF-EFF (cuadros 3.10 y 3.11). Como indicadores del contenido informativo que cada variable contiene específicamente sobre la variable *GASTOH* o *riquezane1*, en cada caso, estos coeficientes serán empleados para ponderarlas de forma diferenciada en el cálculo de las correspondientes medidas de proximidad entre los hogares (los usaremos como valores para los parámetros W_k que aparecen en la distancia de Gower).

⁵⁶ Estos factores de ponderación toman el valor 3 para las variables de ingresos y gastos totales de los hogares, el valor 2 para las componentes de la *renta disponible*, y el valor 1 para el resto de variables.

Cuadro 3.8 Variables matching para el proceso de fusión estadística BDH base 2013 año 2009 vs. EPF año 2008

	Variables	Factor de ponderación en cada tipo de hogar										
		W ₀₁	W ₀₂	W ₀₃	W ₀₇	W ₀₈	W ₀₉	W ₁₀	W ₁₁	W ₁₂	W ₁₃	W ₁₄
SUSTENTADOR PRINCIPAL	sexo_SP	0	0	0	106	0	0	0	0	0	0	0
	edad_SP	767	246	206	183	204	259	368	228	267	347	271
	gedad_SP	767	246	206	183	204	259	368	228	267	347	271
	ned_SP	0	222	0	0	188	0	255	167	209	0	0
	naci_SP	0	157	214	87	203	126	0	0	205	228	0
	cse_SP	0	0	196	0	0	0	0	0	0	0	0
	estcivil_SP	0	115	148	73	94	91	196	83	87	177	115
	situprof_SP	371	204	192	123	111	182	0	0	0	0	0
	tipcont_SP	0	0	193	141	0	0	0	0	0	0	0
	ocupac_SP	0	0	194	94	0	0	0	0	0	0	0
2º ADULTO	FP_ing2_m	989	411	425	266	232	325	471	342	295	375	270
	sexo_2A	0	0	0	0	89	0	0	0	0	0	0
	edad_2A	0	0	0	109	94	103	231	0	0	0	0
	gedad_2A	0	0	0	109	94	103	231	0	0	0	0
	estcivil_2A	0	0	0	83	81	82	217	0	0	0	0
	ned_2A	0	0	0	0	0	0	216	0	0	0	0
	naci_2A	0	0	0	88	198	123	226	0	0	0	0
RESUMEN DE COMPOSICIÓN DEL HOGAR	n_hombre	0	0	0	89	115	74	0	0	75	0	0
	n_mujer	0	0	0	89	115	78	0	0	75	0	118
	n_soltero	0	130	0	96	118	94	0	100	107	163	110
	n_casado	0	0	221	0	107	81	204	113	100	181	0
	n_separado	0	0	0	0	85	74	0	0	94	177	95
	n_viudo	0	0	0	0	92	0	212	0	0	195	0
	n_divorciado	440	173	162	103	77	0	232	0	0	0	97
	n_espanola	0	163	241	103	198	0	0	162	175	166	0
	n_exUE24	0	0	249	96	85	71	190	0	76	0	0
	n_exResE	0	189	171	0	219	99	0	149	0	0	0
	n_sinestud	0	148	0	0	0	0	0	110	86	0	0
	n_eduprimaria	0	280	235	240	173	137	159	160	122	0	106
	n_edusec1	0	175	0	0	132	0	0	96	125	157	0
	n_edusec2	0	130	273	149	82	92	0	77	86	162	109
	n_edusuperior	365	348	404	338	256	181	287	189	223	251	179
	n_ocupado	0	298	186	129	182	179	254	200	200	0	0
	n_desempleado	0	208	0	0	189	0	223	0	153	0	0
	n_estudiante	0	0	0	0	104	0	229	124	136	160	121
	n_retirado	0	0	138	0	82	109	0	82	0	0	0
	n_labhogar	0	0	149	98	133	0	0	133	111	175	0
	n_otrinactivo	0	0	0	0	0	0	0	106	0	0	0
HOGAR (variables sociodemográficas)	nmh	0	0	0	0	0	221	310	0	0	259	290
	nmh16	0	0	0	0	0	221	440	295	325	299	338
	nmh14	0	0	0	0	0	221	451	284	317	306	312
	REGION_m	0	288	288	281	208	187	0	239	216	0	251
	HAB_m	576	227	399	387	195	237	201	241	226	374	163
	TIPOVIV_m	0	0	169	163	101	102	0	91	89	178	118
	REGTENEVIV_m	0	0	0	93	0	116	0	0	0	243	192
	NUMHABIT_m	0	125	111	114	99	98	0	136	146	318	166
HOGAR (variables económicas)	n_percep	0	0	0	90	0	0	0	0	0	0	0
	IIC10_INTOTALES_m	1222	778	832	737	677	669	982	739	727	928	720
	PI_INTOTALES_m	1222	778	832	737	677	669	982	739	727	928	720
	IND1_IT_m	1222	778	832	737	677	669	982	739	727	928	720
	IIC10_IND1_IT_m	1222	778	832	737	677	669	982	739	727	928	720
	I_CPRO	419	206	0	0	0	0	0	143	0	225	147
	I_CAJE	0	233	246	0	136	0	293	0	125	192	0
	I_PENS	0	186	182	95	86	174	0	98	0	0	0
	I_OTRI	394	0	0	0	0	0	0	0	0	0	0

Fuente: Elaboración propia a partir de la ECV base 2013 año 2009 y la EPF año 2008.

Cuadro 3.9 Variables *matching* para el proceso de fusión estadística BDH base 2013 año 2012 vs. EPF año 2011

	Variables	Factor de ponderación en cada tipo de hogar										
		W ₀₁	W ₀₂	W ₀₃	W ₀₇	W ₀₈	W ₀₉	W ₁₀	W ₁₁	W ₁₂	W ₁₃	W ₁₄
SUSTENTADOR PRINCIPAL	sexo_SP	299	0	0	88	0	82	0	0	0	0	133
	edad_SP	794	171	203	186	138	239	353	260	268	315	274
	gedad_SP	794	171	203	186	138	239	353	260	268	315	274
	ned_SP	0	0	0	0	0	0	0	199	0	0	0
	naci_SP	0	170	0	92	193	140	0	237	233	371	220
	estcivil_SP	0	0	111	91	92	107	202	102	102	0	139
	situprof_SP	0	184	0	107	0	137	0	0	229	358	0
	tipcont_SP	0	0	0	101	0	0	0	0	0	0	0
	activesb1_SP_m	0	169	0	87	131	115	0	0	0	0	156
2º ADULTO	sexo_2A	0	0	0	89	0	118	0	0	94	0	212
	edad_2A	0	0	0	91	82	89	186	0	0	0	0
	gedad_2A	0	0	0	91	82	89	186	0	0	0	0
	estcivil_2A	0	0	0	81	90	113	185	0	0	0	0
	ned_2A	0	0	0	0	0	0	179	0	0	0	0
	naci_2A	0	0	0	97	179	141	0	0	0	0	219
RESUMEN DE COMPOSICIÓN DEL HOGAR	n_hombre	299	0	0	0	0	76	0	0	0	0	107
	n_mujer	299	0	0	0	0	79	162	0	0	154	88
	n_soltero	0	0	0	78	104	72	160	118	95	0	96
	n_casado	0	0	152	90	109	96	220	124	108	0	121
	n_separado	0	0	125	0	100	97	185	0	78	0	0
	n_viudo	0	0	0	75	80	80	182	0	0	0	0
	n_divorciado	0	0	0	85	0	97	227	80	78	0	122
	n_espanola	301	186	0	82	200	113	0	211	190	0	151
	n_exUE24	0	0	0	81	133	86	0	122	86	0	113
	n_exResE	0	213	115	102	187	95	179	192	0	0	147
	n_sinestud	0	0	0	0	87	0	0	144	166	318	0
	n_eduprimaria	0	233	325	0	0	0	0	161	0	0	0
	n_edusec1	0	197	0	0	0	0	178	154	160	0	0
	n_edusec2	315	130	213	128	0	98	175	91	81	0	123
	n_eduperior	310	309	363	310	268	169	265	235	233	0	206
	n_ocupado	0	326	0	136	234	173	221	259	290	471	232
	n_desempleado	341	357	0	110	248	132	219	251	252	368	185
	n_estudiante	0	0	0	0	0	105	188	106	129	173	158
	n_retirado	0	0	0	0	0	68	0	0	0	0	99
	n_labhogar	0	0	118	103	90	86	0	123	118	0	0
HOGAR (variables sociodemográficas)	nmh	0	0	0	0	0	193	253	0	0	288	208
	nmh16	0	0	0	0	0	193	359	266	305	354	259
	nmh14	0	0	0	0	0	193	359	241	312	359	252
	REGION_m	0	288	314	244	252	218	0	231	244	0	0
	HAB_m	736	291	352	331	188	204	317	196	199	424	170
	TIPOVIV_m	0	135	163	155	93	97	168	118	88	0	127
	REGTENEVIV_m	0	0	0	0	0	179	0	0	0	0	207
	NUMHABIT_m	0	0	0	0	121	0	0	145	0	0	0
HOGAR (variables económicas)	n_percep	398	237	0	98	0	0	0	0	0	0	0
	FP_ing2_m	582	407	205	235	322	316	414	459	470	699	392
	IIC10_INTOTALES_m	996	809	722	694	728	700	880	850	882	1031	823
	Pi_INTOTALES_m	996	809	722	694	728	700	880	850	882	1031	823
	IND1_IT_m	996	809	722	694	728	700	880	850	882	1031	823
	IIC10_IND1_IT_m	996	809	722	694	728	700	880	850	882	1031	823
	I_CPRO	0	151	0	0	85	0	0	141	116	256	174
	I_CAJE	0	268	0	0	235	0	0	0	267	0	0
	I_PENS	0	0	0	0	0	103	220	90	0	0	0
	I_OTRI	0	125	0	0	0	0	0	0	0	0	0

Fuente: Elaboración propia a partir de la ECV base 2013 año 2012 y la EPF año 2011.

Cuadro 3.10 Variables matching para el proceso de fusión estadística BDH base 2013 año 2009 - EPF año 2008 vs. EF2008

		Factor de ponderación en cada tipo de hogar											
		Variables	W ₀₁	W ₀₂	W ₀₃	W ₀₇	W ₀₈	W ₀₉	W ₁₀	W ₁₁	W ₁₂	W ₁₃	W ₁₄
SUSTENTADOR PRINCIPAL	sexo_SP	476	0	190	93	184	169	306	134	223	532	335	
	edad_SP	1073	448	0	260	546	418	677	476	464	726	660	
	gedad_SP	1073	448	0	260	546	418	677	476	464	726	660	
	ned_SP	0	0	0	0	163	0	0	0	283	0	0	
	estcivil_SP	0	0	123	186	0	193	0	0	0	0	0	
	situac2_SP	0	0	0	0	0	0	0	0	0	0	0	
2º ADULTO	sexo_2A	0	0	0	200	177	249	404	159	221	532	266	
	edad_2A	0	0	0	0	0	147	0	0	0	0	257	
	gedad_2A	0	0	0	0	0	147	0	0	0	0	257	
	estcivil_2A	0	0	0	183	0	157	0	0	0	0	0	
	ned_2A	0	0	0	0	0	219	0	187	0	0	0	
	ocupac_2A	0	0	0	0	0	0	0	0	0	0	0	
RESUMEN DE COMPOSICIÓN DEL HOGAR	nmh	0	0	0	0	0	316	521	0	0	480	532	
	nmh16	0	0	0	0	0	316	547	482	363	705	404	
	nmh14	0	0	0	0	0	316	456	516	399	604	395	
	REGTENEVIV3_m	0	0	421	443	0	0	0	0	335	0	0	
	n_dep	0	0	0	0	0	0	0	0	0	0	0	
HOGAR (variables sociodemográficas)	n_hombre	476	0	190	0	154	203	278	158	168	0	0	
	n_mujer	476	0	190	0	154	179	0	158	168	0	0	
	n_soltero	0	233	123	143	0	0	0	0	141	0	0	
	n_casado	0	0	109	126	0	173	0	0	0	0	0	
	n_separado	0	194	104	135	136	182	295	133	161	196	146	
	n_viuado	0	309	177	129	179	177	370	117	91	0	0	
	n_divorciado	0	256	141	109	151	112	0	124	131	0	257	
	n_sinestud	0	176	230	0	116	0	0	123	377	0	0	
	n_eduprimaria	0	0	99	175	0	188	0	173	206	0	0	
	n_edusec1	438	227	191	173	202	140	267	215	0	0	0	
	n_edusec2	0	0	252	205	124	187	281	168	0	0	0	
	n_eduperior	0	360	295	273	206	0	0	180	0	0	0	
	n_ocupado	0	0	145	179	161	175	0	114	199	0	0	
	n_desempleado	0	345	0	106	223	216	241	184	290	0	0	
	n_estudiante	0	0	0	93	200	167	281	0	185	0	0	
	n_refirado	0	0	0	80	172	107	0	148	138	440	190	
	n_labhogar	0	126	0	108	246	149	0	119	156	0	0	
n_otrinactivo	0	0	0	0	0	0	0	0	204	0	0		
HOGAR (variables económicas)	n_percep	0	133	0	182	0	0	0	0	0	0	0	
	IIC10_RT_B	2196	807	665	711	707	698	958	707	733	1309	815	
	I_CPRO	476	218	135	0	319	348	355	284	376	587	0	
	I_CAJE	0	0	189	184	0	0	0	0	0	0	154	
	I_PENS	0	160	95	122	233	220	258	131	158	377	261	
	I_DES	271	301	0	0	0	0	0	0	0	0	0	
	I_OSPS	0	232	109	143	167	137	254	238	287	510	193	
	I_RCYP	0	443	0	0	0	0	364	438	0	424	0	
	I_OTRI	0	0	0	130	238	0	231	208	292	513	314	
	FP_ING_m	732	539	399	309	408	463	584	387	495	816	394	
	IIC10_RMCP_B	0	0	0	75	123	0	169	123	169	0	152	
	PI_RMCP_B	0	0	0	75	123	0	169	123	169	0	152	
	IIC10_RNMCP_B	0	86	0	46	62	65	0	0	74	0	79	
	PI_RNMCP_B	0	86	0	46	62	65	0	0	74	0	79	
	IIC10_RMIX_B	789	255	121	158	249	272	349	246	315	534	297	
	PI_RMIX_B	789	255	121	158	249	272	349	246	315	534	297	
	IIC10_RMCA_B	0	0	0	0	149	0	0	0	0	0	0	
	PI_RMCA_B	0	0	0	0	149	0	0	0	0	0	0	
	IIC10_RNMCA_B	0	112	0	63	82	59	144	86	75	0	0	
	PI_RNMCA_B	0	112	0	63	82	59	144	86	75	0	0	
	IIC10_RASAL_B_m	893	427	162	245	307	264	606	380	371	668	426	
	PI_RASAL_B_m	893	427	162	245	307	264	606	380	371	668	426	
	IIC10_IMCAP_B	0	0	0	185	190	0	0	187	173	0	242	
	PI_IMCAP_B	0	0	0	185	190	0	0	187	173	0	242	
	IIC10_IMPRO_B	0	148	84	106	116	94	0	86	102	0	125	
	PI_IMPRO_B	0	148	84	106	116	94	0	86	102	0	125	
	IIC10_RPRO_B	0	0	0	208	194	0	0	189	187	0	257	
	PI_RPRO_B	0	0	0	208	194	0	0	189	187	0	257	
	IIC10_RPRIH_B_m	0	0	0	0	0	0	0	0	0	0	0	
	PI_RPRIH_B_m	0	0	0	0	0	0	0	0	0	0	0	
	IIC10_PSDS_B	0	0	0	0	96	0	0	97	110	0	0	
	PI_PSDS_B	0	0	0	0	96	0	0	97	110	0	0	
	IIC10_PSIJUB_B	0	0	0	0	0	126	0	0	0	0	0	
	PI_PSIJUB_B	0	0	0	0	0	126	0	0	0	0	0	
	IIC10_PSSUP_B	0	0	0	0	0	81	0	0	0	0	0	
	PI_PSSUP_B	0	0	0	0	0	81	0	0	0	0	0	
	IIC10_PSIJYS_B_m	0	127	0	0	0	152	0	99	83	0	0	
	PI_PSIJYS_B_m	0	127	0	0	0	152	0	99	83	0	0	
	IIC10_PSENF_B	0	92	0	54	0	66	0	0	73	0	84	
	PI_PSENF_B	0	92	0	54	0	66	0	0	73	0	84	
	IIC10_PSIINV_B	0	102	0	0	78	67	0	86	0	0	96	
	PI_PSIINV_B	0	102	0	0	78	67	0	86	0	0	96	
	IIC10_PSEEL_B_m	0	93	0	0	76	66	0	92	0	0	101	
	PI_PSEEL_B_m	0	93	0	0	76	66	0	92	0	0	101	
	IIC10_PSEST_B	0	0	0	57	67	53	0	89	95	0	107	
	PI_PSEST_B	0	0	0	57	67	53	0	89	95	0	107	
	IIC10_PSFYV_B_m	0	88	0	48	61	60	0	0	80	0	94	
	PI_PSFYV_B_m	0	88	0	48	61	60	0	0	80	0	94	
	IIC10_RPEPP_B	0	86	0	54	0	73	0	0	80	0	0	
	PI_RPEPP_B	0	86	0	54	0	73	0	0	80	0	0	
IIC10_PRESOC_B_m	778	322	326	343	276	308	473	341	383	679	333		
PI_PRESOC_B_m	778	322	326	343	276	308	473	341	383	679	333		
IIC10_TPRDOH_B	0	99	86	55	71	73	0	84	85	0	106		
PI_TPRDOH_B	0	99	86	55	71	73	0	84	85	0	106		
IIC10_TCPER_B_m	0	0	0	0	0	0	0	0	0	0	0		
PI_TCPER_B_m	0	0	0	0	0	0	0	0	0	0	0		
PI_RT_B	2196	807	665	711	707	698	958	707	733	1309	815		
PI_GASTOH	1679	808	610	646	580	649	1076	591	596	1236	803		
IIC10_GASTOH	1679	808	610	646	580	649	1076	591	596	1236	803		
IIC10_GCF_01	1084	767	597	505	506	567	1055	675	595	1264	706		
PI_GCF_01	1084	767	597	505	506	567	1055	675	595	1264	706		

Fuente: Elaboración propia a partir de la ECV base 2013 año 2009 y las EPF y EFF año 2008.

Cuadro 3.11 Variables matching para el proceso de fusión estadística BDH base 2013 año 2012 - EPF año 2011 vs. EF2011

		Factor de ponderación en cada tipo de hogar											
Variables		W ₀₁	W ₀₂	W ₀₃	W ₀₇	W ₀₈	W ₀₉	W ₁₀	W ₁₁	W ₁₂	W ₁₃	W ₁₄	
SUSTENTADOR PRINCIPAL	sexo_SP	0	231	143	130	141	143	300	220	125	514	205	
	edad_SP	1322	549	0	295	661	430	451	500	359	845	572	
	edad_SP	1322	549	0	295	661	430	451	500	359	845	572	
	ned_SP	0	0	299	0	0	0	0	203	0	0	0	
	estcivil_SP	0	0	155	0	0	0	0	0	143	0	0	
	situprof_SP	0	0	153	165	244	223	0	0	0	0	0	
	tipcont_SP	0	236	140	166	235	224	337	222	257	583	304	
2º ADULTO	sexo_2A	0	0	0	155	216	154	376	215	122	499	308	
	edad_2A	0	0	0	0	0	196	0	229	0	0	0	
	edad_2A	0	0	0	0	0	196	0	229	0	0	0	
	ocupac_2A	0	0	0	127	0	182	0	0	0	0	0	
	situprof_2A	0	0	0	142	0	0	0	0	0	0	0	
	situprof_2A	0	0	0	0	182	226	0	202	211	513	369	
	tipcont_2A	0	0	0	0	0	0	0	0	0	0	0	
RESUMEN DE COMPOSICIÓN DEL HOGAR	nmh	0	0	0	0	0	350	467	0	0	877	461	
	nmh16	0	0	0	0	0	350	640	483	410	773	532	
	nmh14	0	0	0	0	0	350	585	535	332	795	543	
	REGTENEVIV3_m	0	0	441	406	0	0	0	397	447	0	0	
HOGAR (variables sociodemográficas)	n_hombre	0	231	143	179	165	129	0	115	149	0	229	
	n_mujer	0	231	143	179	165	0	0	115	149	0	0	
	n_soltero	0	188	149	121	0	0	314	213	181	0	0	
	n_casado	0	0	145	127	0	0	0	0	136	390	0	
	n_separado	0	170	141	93	114	117	260	219	159	0	0	
	n_viudo	0	258	176	111	163	157	315	117	95	292	155	
	n_divorciado	0	233	177	112	178	121	373	102	94	264	0	
	n_sinestud	0	178	295	0	151	0	270	94	289	0	0	
	n_eduprimaria	0	0	264	0	0	0	0	0	0	0	0	
	n_edusec1	864	230	270	0	236	0	283	0	0	0	0	
	n_edusec2	672	293	176	163	186	185	301	202	0	416	0	
	n_edusuperior	712	262	354	0	227	0	0	0	189	361	0	
	n_ocupado	398	189	202	191	188	0	314	0	297	0	0	
	n_desempleado	0	174	147	151	261	249	259	212	0	0	347	
	n_estudiante	0	0	0	143	213	0	308	187	176	0	0	
	n_retirado	0	0	0	152	0	126	254	109	94	0	158	
	n_labhogar	0	217	0	91	260	0	351	189	167	0	252	
	n_otrinactivo	0	0	202	0	0	0	0	0	0	0	0	
	HOGAR (variables económicas)	n_percep	0	158	0	0	153	0	0	0	0	0	0
I_CPRO		0	0	160	0	295	345	417	347	426	533	452	
I_CAJE		0	0	117	0	0	0	0	0	0	647	0	
I_PENS		0	331	0	0	301	0	0	188	0	200	0	
I_OSFS		0	254	141	201	107	202	328	215	184	0	240	
I_RCYP		839	424	0	460	0	0	332	447	422	429	0	
I_OTRI		0	0	0	118	0	0	0	138	0	0	399	
FP_ING_m		739	437	346	293	437	428	609	397	461	930	474	
IIC10_RMCP_B		0	0	72	0	117	147	0	0	159	0	181	
PI_RMCP_B		0	0	72	0	117	147	0	0	159	0	181	
IIC10_RNMCP_B		0	0	0	55	0	69	0	79	72	0	123	
PI_RNMCP_B		0	0	0	55	0	69	0	79	72	0	123	
IIC10_RMIX_B		909	213	142	141	238	303	348	265	322	506	420	
PI_RMIX_B		909	213	142	141	238	303	348	265	322	506	420	
IIC10_RMCA_B		0	197	0	0	182	0	0	0	197	0	0	
PI_RMCA_B		0	197	0	0	182	0	0	0	197	0	0	
IIC10_RNMCA_B		0	0	0	49	96	71	0	89	95	0	118	
PI_RNMCA_B		0	0	0	49	96	71	0	89	95	0	118	
IIC10_RASAL_B_m		1074	404	194	217	362	355	519	311	391	701	425	
PI_RASAL_B_m		1074	404	194	217	362	355	519	311	391	701	425	
IIC10_IMCAP_B		0	0	0	208	205	0	0	180	183	0	227	
PI_IMCAP_B		0	0	0	208	205	0	0	180	183	0	227	
IIC10_IMPRO_B		0	162	100	114	84	111	0	110	134	286	123	
PI_IMPRO_B		0	162	100	114	84	111	0	110	134	286	123	
IIC10_RPRO_B		0	265	221	211	190	0	0	0	220	0	250	
PI_RPRO_B		0	265	221	211	190	0	0	0	220	0	250	
IIC10_RPRIH_B_m		668	251	206	208	211	220	284	221	247	443	286	
PI_RPRIH_B_m		668	251	206	208	211	220	284	221	247	443	286	
IIC10_PSDES_B		0	0	0	0	130	0	0	0	128	0	0	
PI_PSDES_B		0	0	0	0	130	0	0	0	128	0	0	
IIC10_PSJUB_B		0	0	0	148	131	135	163	119	0	0	0	
PI_PSJUB_B		0	0	0	148	131	135	163	119	0	0	0	
IIC10_PSSUP_B		0	0	0	58	0	68	0	0	0	0	91	
PI_PSSUP_B		0	0	0	58	0	68	0	0	0	0	91	
IIC10_PSJYS_B_m		0	0	0	0	0	141	0	0	0	0	0	
PI_PSJYS_B_m		0	0	0	0	0	141	0	0	0	0	0	
IIC10_PSENF_B		0	0	57	50	0	74	0	81	0	0	90	
PI_PSENF_B		0	0	57	50	0	74	0	81	0	0	90	
IIC10_PSINV_B		0	0	60	69	68	80	0	103	66	0	111	
PI_PSINV_B		0	0	60	69	68	80	0	103	66	0	111	
IIC10_PSEEL_B_m		0	0	0	70	69	83	0	100	0	0	0	
PI_PSEEL_B_m		0	0	0	70	69	83	0	100	0	0	0	
IIC10_PSEST_B		0	0	0	54	0	61	0	91	93	0	101	
PI_PSEST_B		0	0	0	54	0	61	0	91	93	0	101	
IIC10_PSFYV_B_m		0	0	58	61	0	73	0	82	88	0	107	
PI_PSFYV_B_m		0	0	58	61	0	73	0	82	88	0	107	
IIC10_RPEPP_B		0	0	0	0	0	79	0	75	0	0	93	
PI_RPEPP_B		0	0	0	0	0	79	0	75	0	0	93	
IIC10_PRESOC_B_m		149	392	369	348	341	284	379	302	305	637	347	
PI_PRESOC_B_m		149	392	369	348	341	284	379	302	305	637	347	
IIC10_TPRDOH_B		0	0	0	58	78	74	0	87	86	0	109	
PI_TPRDOH_B		0	0	0	58	78	74	0	87	86	0	109	
IIC10_RT_B		1983	850	748	644	755	811	922	663	858	1329	842	
PI_RT_B		1983	850	748	644	755	811	922	663	858	1329	842	
IIC10_GASTOH		2039	737	721	573	685	554	991	601	722	1373	930	
PI_GASTOH		2039	737	721	573	685	554	991	601	722	1373	930	
IIC10_GCF_01		1668	603	622	506	640	580	924	574	609	1080	697	
PI_GCF_01		1668	603	622	506	640	580	924	574	609	1080	697	

Fuente:Elaboración propia a partir de la ECV base 2013 año 2012 y las EPF y EFF año 2011.

3.1.2.5. Aplicación y evaluación de resultados

De acuerdo con lo expuesto en el apartado anterior, para cada uno de los hogares de un determinado tipo en la ECV base 2013 -presentes en la BDH base 2013- hemos calculado, para cada uno de los años objeto de estudio, la disimilitud que presentaba con cada uno de los hogares del mismo tipo que se recogían en la EPF, utilizando para ello la medida de similaridad de Gower, las variables comunes que resultaron significativas para el tipo de hogar y sus correspondientes ponderaciones calculadas (mostradas en los cuadros 3.8 y 3.9).

Una vez resuelto el cálculo de esta medida de proximidad, como mecanismo de enlace de casos para la fusión de las variables no comunes seleccionadas en el apartado 3.1.2.3, se ha empleado el método de los *k* vecinos más próximos. En el caso general, este método determina, para cada hogar de la ECV, los *k* hogares más próximos de entre los del mismo tipo recogidos en la EPF (es decir, los *k* hogares que menor disimilitud calculada presentan con aquél); y, una vez determinados éstos, resume mediante algún criterio la información que proporcionan las variables no comunes a fusionar.

En particular, en la aplicación que hemos realizado de este método para este trabajo, se ha utilizado el hogar vecino más próximo (*k*=1) de entre los disponibles en cada momento para determinar la información a fusionar, añadiendo directamente la información que proporciona éste en las variables no comunes a fusionar.

A partir de la fusión de los conjuntos de datos BDH base 2013-EPF para los años de referencia de los ingresos 2008 y 2011, hemos llevado a cabo, mediante el mismo procedimiento de fusión, el "emparejamiento" de los hogares de la ECV con los de las EFF 2008 y 2011, respectivamente, por medio de las variables comunes y de sus respectivos pesos presentes en los cuadros 3.10 y 3.11.

Como ya hemos avanzado anteriormente, en el Anexo 2 se muestra el programa informático que hemos desarrollado para la realización del proceso de *matching* que se ha descrito en a lo largo del apartado 3 de este capítulo.

Por último, para garantizar la aplicabilidad⁵⁷ de los ficheros fusionados, es necesario analizar los resultados obtenidos a través del método de fusión presentado. Cada una de las etapas del proceso de *matching* anteriormente descritas tiene por objetivo impactar positivamente en la calidad de los resultados. Sin embargo, en términos de coherencia e integración, los resultados obtenidos a través de las fusiones estadísticas realizadas aún deben ser validados en términos de su potencial para proporcionar estimaciones fiables y exactas.

Rässler (2002) estableció cuatro niveles para poder evaluar la validez de un procedimiento de fusión estadística de forma sistemática:

⁵⁷ Se habla aquí de validez y no de eficiencia: no se emplea aquí ningún criterio similar al del mínimo error cuadrático medio tal como es utilizado en otros campos de la estadística; más bien, aquí el objetivo es evaluar los diferentes niveles de reproducción y preservación de las distribuciones y asociaciones originales.

- **Nivel 1: Preservar valores individuales.** Tras la fusión, los valores reales (no-observados) de las variables Y se reproducen de manera exacta en el fichero receptor: $y_i = \hat{y}_i$ para $i = 1, 2, \dots, n_A$, donde y_i son los valores reales (inobservados), e \hat{y}_i son los imputados, siendo n_A el tamaño del fichero receptor (A). Se busca calcular el número de veces que el valor imputado coincide con el "real" para poder calcular una "tasa de aciertos".
- **Nivel 2: Preservar distribuciones conjuntas.** Después de la fusión, la distribución conjunta real (inobservada) de los tres conjuntos de variables X, Y, Z debe reflejarse correctamente en el fichero imputado. Esto es: $f(X, Y, Z) = \hat{f}(X, Y, Z)$ donde f denota la distribución conjunta observada y \hat{f} la obtenida en el fichero imputado.
- **Nivel 3: Preservar la estructura de correlaciones.** Tras el enlace, el fichero fusionado debe preservar la estructura de correlaciones y momentos de orden superior. Esto es: $cov(X, Y, Z) = \hat{cov}(X, Y, Z)$, donde cov denota la matriz de varianzas-covarianzas observada, y \hat{cov} la calculada sobre el fichero fusionado.
- **Nivel 4: Preservar las distribuciones marginales.** Tras el enlace, las distribuciones marginales de las variables observadas en el fichero donante se reproducen correctamente en el fichero imputado. En particular, deben cumplirse: $f(Y) = \hat{f}(Y)$ y $f(Y|Z) = \hat{f}(Y|Z)$, donde f denota las distribuciones marginales observadas y reales y \hat{f} las imputadas.

En la práctica, la mayoría de los métodos tradicionales se centran en el estudio del nivel cuarto. De hecho, solamente este nivel se podrá contrastar directamente. El cumplimiento de este nivel garantiza que las distribuciones marginales observadas en el fichero donante se reproducen correctamente en el fichero receptor. Siempre que se empleen procedimientos robustos y datos de calidad, es relativamente sencillo alcanzar este nivel; por ello, es el requisito mínimo que se le puede pedir a cualquier ejercicio de enlace (Leulescu *et al.*, 2013).

A continuación, analizamos el grado de cumplimiento de este cuarto nivel en los ficheros fusionados en nuestro trabajo basándonos en dos variables principales imputadas (donadas mediante fusión) en los hogares de la ECV base 2013: el gasto total (GASTOH) y la riqueza neta (*riquezaset*). Los resultados (cuadros 3.12 y 3.13, respectivamente) muestran que el método de imputación presentado anteriormente conserva razonablemente bien las distribuciones marginales de las variables antes y después la fusión.

Cuadro 3.12
ESTADÍSTICOS DESCRIPTIVOS DEL GASTO DE LOS HOGARES
OBSERVADO (EPF) E IMPUTADO (BDH B2013)

		Mean	SE	Min	20th. Perc.	Median	80th. Perc.	Max
2008	GASTOH Observado	31.953	19.934	1.761	16.392	27.675	44.483	332.933
	GASTOH Imputado	31.686	19.733	1.761	15.956	27.309	44.460	332.933
2011	GASTOH Observado	29.482	18.457	1.543	15.601	25.423	40.549	730.703
	GASTOH Imputado	29.109	18.220	1.756	15.318	25.048	40.198	730.703

Fuente: Elaboración propia a partir de las ECV base 2013 y las EPF.

Cuadro 3.13
**ESTADÍSTICOS DESCRIPTIVOS DE LA RIQUEZA NETA DE LOS HOGARES
 OBSERVADA (EFF) E IMPUTADA (BDH B2013)**

		Mean	SE	Min	20th. Perc.	Median	80th. Perc.	Max
2008	riquezane _t Observada	283.857	1.184.250	-1.278.787	49.066	177.957	374.324	409.241.674
	riquezane _t Imputada	295.869	969.213	-1.278.787	51.332	181.204	376.857	149.130.000
2011	riquezane _t Observada	266.892	749.394	-57.332.274	41.707	153.100	350.039	156.744.991
	riquezane _t Imputada	281.462	958.785	-57.332.274	41.011	155.000	350.156	124.460.000

Fuente: Elaboración propia a partir de las ECV base 2013 y las EFF.

Como puede verse y excluyendo los valores mínimo y máximo que obviamente son proclives a presentar, en general, mayores diferencias (y aunque no siempre en nuestros casos), la variable *GASTOH* presenta errores relativos para las diferentes medidas de localización presentadas en torno al 1%, presentándose el mayor desajuste para el percentil 20 (2,66%). Por su parte, la variable *riquezane_t*, presenta, en general, errores relativos similares (en torno al 1% para las medianas y percentiles en general), aunque algo mayores comparativamente para el caso del percentil 20 en 2008 (4,62%) o las medias (4,23% en 2008 y 5,46% en 2011). Esto nos hace pensar que una posible causa haya sido la segmentación empleada en ambos casos (la clasificación sociológica de los tipos de hogares). Si bien, parece haberse mostrado adecuada para la fusión de variables relativas al gasto, parece no ser del todo eficiente para la fusión de variables relacionadas con el grado de riqueza. Quizás en este caso, debemos explorar alguna otra segmentación de hogares más adecuada para este tipo de variables; lo que posponemos para futuras investigaciones.

En definitiva, y pese a las posibles limitaciones manifestadas, los resultados obtenidos de sendos procesos de fusión presentados en sus distintas etapas a lo largo de este apartado 3.3, consistentes en dos grandes bases de datos para cada uno de los años de referencia de los ingresos (2008 y 2011): ***BDH base 2013 año 2009-EPF 2008-EFF 2008*** y ***BDH base 2013 año 2012-EPF 2011-EFF 2011***, nos van a permitir disponer de información integrada sobre ingresos, gastos y situación financiera de los hogares y afrontar así la posibilidad de corregir la infradeclaración de rentas detectada en los hogares de la ECV. Para ello, en primer lugar, haremos uso de gran cantidad de información contenida en estas bases de datos para estimar las variables no disponibles en la ECV que ya pusimos de manifiesto en el capítulo II, necesarias para el cálculo de la renta disponible de los hogares de forma homogénea a la Contabilidad Nacional. Una vez dispongamos de esta información faltante, con la finalidad de proceder a la corrección de las distintas partidas de la renta disponible de los hogares de la ECV y a la estimación de una renta congruente con la publicada por la CNE según el marco contable de referencia -SEC 2010-, presentaremos un conjunto de indicadores de incongruencia diseñados en base a las relaciones de la literatura y teoría económicas y a partir del aprovechamiento de variables proxy que nos proporcionan las fuentes fusionadas.

4. Estimación de partidas no disponibles en las fuentes de estudio

Como ya estudiamos en el capítulo II, existen un conjunto de partidas de la CNE que no se recogen en la ECV y que se hace necesario estimar para obtener una comparación homogénea entre ambas fuentes. Asimismo, necesitaremos estimar algunos datos en nuestro marco de referencia, ya que no están disponibles en la CNE Base 2010, ya sea porque debido a las nuevas normas de publicación de la nueva base, ahora se publican con menor nivel de desagregación, o porque dichas cifras no estén retropoladas en Base 2010. Aprovechando la información adicional (EPF y EFF) que hemos fusionado a los hogares de la ECV en nuestra BDH, dedicaremos este apartado a la estimación de aquellas variables desconocidas.

4.1. Estimación de partidas no disponibles en la CNE Base 2010

A continuación, expondremos, por orden de aparición en las Cuentas de Renta del sector hogares (ver cuadros 2.7 y 2.8, en el capítulo II), la serie de variables que procederemos a estimar en la CNE Base 2010, tomando como referencia los resultados que dichas partidas obtuvieron en la CNE Base 2008.

Comenzamos por la estimación de algunos de los componentes del EBE de la rama de Alquiler Imputado para el año 2008 en Base 2010. Dichos resultados están disponibles en Base 2008 en la Tabla de Origen y Destino de la economía española para dicho año, pero no lo están en su totalidad en Base 2010. Estaríamos refiriéndonos a los Consumos Intermedios relativos a los productos siguientes de dicha rama: 'Servicios de intermediación financiera', que designaremos como $AIVPP_CI_IPPH_Cne$, 'Trabajos de construcción especializados' ($AIVPP_CI_TCE_Cne$) y 'Servicios de seguros y planes de pensiones' ($AIVPP_CI_Seg_Cne$). Dado que conocemos el peso de cada uno de estos productos sobre el total de Consumos Intermedios de la rama Alquiler Imputado para el año 2008, publicados en la TOD de dicho año en Base 2008, para su estimación en Base 2010 haremos uso de ellos multiplicándolos por el total de Consumos Intermedios de dicha rama en Base 2010 que sí ha publicado el INE. De este modo, tendremos:

$$AIVPP_CI_IPPH_Cne_{B2010}^{2008} = \frac{AIVPP_CI_IPPH_Cne_{B2008}^{2008}}{AIVPP_CI_Cne_{B2008}^{2008}} * AIVPP_CI_Cne_{B2010}^{2008}$$

$$AIVPP_CI_TCE_Cne_{B2010}^{2008} = \frac{AIVPP_CI_TCE_Cne_{B2008}^{2008}}{AIVPP_CI_Cne_{B2008}^{2008}} * AIVPP_CI_Cne_{B2010}^{2008}$$

$$AIVPP_CI_Seg_Cne_{B2010}^{2008} = \frac{AIVPP_CI_Seg_Cne_{B2008}^{2008}}{AIVPP_CI_Cne_{B2008}^{2008}} * AIVPP_CI_Cne_{B2010}^{2008}$$

Dichos componentes sí se conocen para el año 2011 en Base 2010, ya que la TOD de dicho año sí está publicada en Base 2010, por lo que sólo será necesario estimarlos para el año 2008.

Otra información que no se encuentra disponible en Base 2010 para los años 2008 y 2011 en las Cuentas No Financieras del Sector Hogares publicadas por el INE es el desglose de las Cotizaciones sociales a cargo de los empleadores ($D.12_r_Cne$) en Cotizaciones sociales efectivas ($D.121_r_Cne$) y Cotizaciones sociales imputadas ($D.122_r_Cne$). Para su estimación haremos uso nuevamente de los datos publicados sobre estas partidas en Base 2008, teniendo que:

$$D.121_r_Cne_{B2010}^{2008,2011} = \frac{D.121_r_Cne_{B2008}^{2008,2011}}{D.12_r_Cne_{B2008}^{2008,2011}} * D.12_r_Cne_{B2010}^{2008,2011}$$

$$D.122_r_Cne_{B2010}^{2008,2011} = \frac{D.122_r_Cne_{B2008}^{2008,2011}}{D.12_r_Cne_{B2008}^{2008,2011}} * D.12_r_Cne_{B2010}^{2008,2011}$$

Por otra parte, del total de empleos de las Cotizaciones Sociales Netas de CNE ($D.61_e_Cne$) no conocemos su desglose en CNE Base 2010 en 2008 y en 2011, pero sabemos que las Cotizaciones sociales efectivas a cargo de los empleadores ($D.611_e_Cne$) corresponden al flujo $D.121_r_Cne$, estimado anteriormente, y las Cotizaciones sociales imputadas a cargo de los empleadores ($D.612_e_Cne$) representan la contrapartida de las prestaciones sociales directas de los empleadores pagadas directamente por los empleadores a sus asalariados, antiguos asalariados y otros derechohabientes, correspondiendo por tanto al flujo $D.122_r_Cne$.

4.2. Estimación de partidas no disponibles en la ECV

Este apartado lo dedicaremos a detallar las soluciones abordadas para la estimación de aquellas partidas no recogidas en la ECV cuya disponibilidad resulta primordial para llevar a cabo una comparación homogénea respecto a la CNE. Para ello, seguiremos el orden de aparición en las Cuentas de Renta del sector hogares (ver cuadros 2.7 y 2.8).

Comenzaremos por los componentes del EBE de la rama de Alquiler Imputado. La ECV recoge razonablemente bien y de forma conjunta la Producción correspondiente a dicha rama menos los consumos intermedios asociados a las mejoras llevadas a cabo en la vivienda principal ($AIVPP_P - AIVPP_CI_TCE$) -con coberturas del 108,4% en 2008 y 94% en 2011, pero, sin embargo, no ocurre lo mismo con el resto de los Consumos Intermedios (los correspondientes a los Servicios de Intermediación Financiera, $AIVPP_CI_IPPH$, alcanzan un valor aproximadamente tres veces superior a los de CNE y $AIVPP_Seg_IPPH$ no se recogen) ni con los Otros impuestos netos sobre la producción asociados a dicha rama. Por tanto, dado $AIVPP_P - AIVPP_CI_TCE$ para el hogar, utilizaremos los pesos de los componentes de los Consumos Intermedios sobre la Producción de Alquiler Imputado en CNE para estimar dichos componentes no disponibles en la ECV. Así:

$$AIVPP_CI_Seg = \frac{AIVPP_CI_Seg_Cne}{(AIVPP_P_Cne - AIVPP_TCE_Cne)} * (AIVPP_P - AIVPP_CI_TCE)$$

$$AIVPP_CI_OIN = \frac{AIVPP_CI_OIN_Cne}{(AIVPP_P_Cne - AIVPP_TCE_Cne)} * (AIVPP_P - AIVPP_CI_TCE)$$

De forma análoga procederemos para la estimación en la ECV del Consumo de Capital Fijo asociado al EBE:

$$CCFEE = \frac{CCFEE_Cne}{(AIVPP_P_Cne - AIVPP_TCE_Cne)} * (AIVPP_P - AIVPP_CI_TCE)$$

Por su parte, para buscar una solución a la sobreestimación de la variable *AIVPP_CI_IPPH*, vamos a estimarla con la información que nos proporciona la EFF bajo el nombre de *AIVPP_CI_IPPH_Est*. Para ello, en primer lugar, la EFF nos proporciona información sobre el total de Intereses asociados al Préstamo/s Hipotecario vinculado a la compra de la vivienda principal (*IPPH_B_EFF*) a partir de la información relacionada con los mismos ($p2_9_i=1$, siendo $p2_9_i$ el tipo de préstamo i con el que se financió la compra de la vivienda principal ($i = 1,2,3,4$) y "1" la modalidad de esta variable que se utiliza para designar el préstamo/s hipotecario), de forma que:

$$IPPH_B_EFF = \sum_{i=1}^4 p2_12_i_eff * \left(\frac{p2_13_i_eff}{100} \right)$$

siendo,

$p2_12_i_eff$: Capital pendiente de amortizar del préstamo i -sin contar los intereses- con el que se financió la compra de la vivienda principal ($i = 1,2,3,4$);

$p2_13_i_eff$: Tipo de interés (%) del préstamo i con el que se financió la compra de la vivienda principal ($i = 1,2,3,4$).

Por otra parte, de la EFF podemos conocer también el montante anual (amortización e intereses) que destina el hogar al pago del préstamo/s vinculado a la compra de la vivienda principal (*PagosAVP*):

$$PagosAVP = 12 * \sum_{i=1}^4 p2_18_i_eff$$

siendo,

$p2_18_i_eff$: Importe mensual que paga el hogar en la actualidad por el préstamo i -incluyendo amortización e intereses- con el que se financió la compra de la vivienda principal ($i = 1,2,3,4$).

Para conocer cuál es la proporción de intereses hipotecarios pagados en un año calculamos la variable auxiliar *KF1* como:

$$KF1 = \max\left(0, \min\left(1, \frac{IPPH_B_EFF}{PagosAVP}\right)\right)$$

Finalmente obtendremos $AIVPP_CI_IPPH_Est$ como:

$$AIVPP_CI_IPPH_Est = IPPH_B_EFF * KF1$$

Otra de las variables que necesitamos estimar para contar con una aproximación más cercana a la CNE del concepto de Renta Mixta que recoge la ECV son los ingresos que obtiene el hogar por el alquiler de partes de su vivienda principal o de otras propiedades inmobiliarias distintas de la vivienda principal. A esta variable la denominaremos $IMALQPRO_B$. La ECV sí recoge este tipo de rentas, pero lo hace junto a aquellas recibidas por el alquiler de terrenos en la variable $HY040G$ (a la que denominaremos $IMPRO_B$), razón por la cual necesitamos buscar una fuente alternativa que nos permita estimar la parte de ingresos destinada al alquiler de propiedades y la destinada al alquiler de terrenos (esta parte, a su vez, nos servirá para estimar la partida *Rentas de la tierra* de CNE –cuya rúbrica es $D_45_re_Cne$ – para cuya estimación en la ECV utilizaremos la variable $IMALQTER_B$. Así, para estimar $IMALQPRO_B$ nos apoyaremos en la información que hemos fusionado en nuestra BDH base 2013 sobre la EFF, en concreto en las variables que nos proporcionan datos de ingresos y pagos del hogar relacionados con el alquiler de habitaciones de la vivienda principal en propiedad y/o de propiedades inmobiliarias que no son una finca o solar –partes de una vivienda, vivienda, oficina, local...–). De forma que:

$$IMALQPRO_B = IMPRO_B - IMALQTER_B$$

siendo

$$IMALQTER_B = \min\left\{IMALQTER_B_EFF, \left[IMPRO_B \frac{IMALQTER_B_EFF}{(IMALQPRO_B_EFF + IMALQTER_B_EFF)}\right]\right\}$$

donde,

$$IMALQPRO_B_EFF = \sum_{j=1}^4 \left\{ 12 * [(p2_43_i_eff - p2_48_i_eff) * I(p2_35a_i_eff \neq 3)] - \sum_{j=1}^4 \left\{ 12 * \left[\left(p2_61_4_eff + \sum_{j=1}^3 p2_61_j_i_eff \right) * I(p2_35a_i_eff \neq 3) \right] \right\} \right\}$$

$$IMALQTER_B_EFF = \sum_{j=1}^4 \left\{ 12 * [(p2_43_i_eff - p2_48_i_eff) * I(p2_35a_i_eff = 3)] - \sum_{j=1}^4 \left\{ 12 * \left[\left(p2_61_4_eff + \sum_{j=1}^3 p2_61_j_i_eff \right) * I(p2_35a_i_eff = 3) \right] \right\} \right\}$$

siendo

$p2_{43_i_eff}$: Ingresos mensuales del hogar por el alquiler de la propiedad inmobiliaria i –excluyendo la vivienda principal- ($i = 1,2,3,4$);

$p2_{48_i_eff}$: Reformas o inversiones llevadas a cabo por el hogar sobre la propiedad inmobiliaria i alquilada, distinta de la vivienda principal ($i = 1,2,3,4$);

$p2_{61_j_i_eff}$: Importe mensual que paga el hogar en la actualidad por el préstamo j que tiene contraído sobre la propiedad inmobiliaria i alquilada, distinta de la vivienda principal ($i = 1,2,3,4 ; j = 1,2,3$);

$p2_{35a_i_eff} \neq 3$: Tipo de propiedad inmobiliaria i distinta de una finca o solar que posee el hogar ($i = 1,2,3,4$);

$I(.)$: Función indicador.

Para estimar el Consumo de Capital Fijo asociado a la Renta Mixta de los hogares en la ECV, $CCFRM$, utilizaremos el peso del Consumo de Capital Fijo de la Renta Mixta sobre la Renta Mixta Neta de CNE y lo multiplicaremos por la Renta Mixta Neta de la BDH (base 2013), para cada uno de los años de estudio. Así, una vez estimada $IMALQPRO_B$, se tendrá que:

$$CCFRM = \frac{CCFRM_Cne}{B_3n_Cne} * (B_3n_Bdh)$$

siendo,

$$B_3n_Bdh = RMCP_B + RNMCP_B + CSEFCP + IMALQPRO_B$$

Por su parte, las Rentas brutas del trabajo percibidas por los menores de 16 años que se recogen en la ECV junto al resto de rentas recibidas por estas unidades, designadas mediante la rúbrica $ORNC1_B$, no ha sido posible estimarlas ya que ni la información de la ECV ni de las fuentes fusionadas (EPF y EFF) ni de la propia CNE, contienen información al respecto.

Otra de las partidas de CNE que no se recogen en la ECV son las *Cotizaciones sociales imputadas a cargo de los empleadores*, que estimaremos a través de la variable $CSIMEMP$, como:

$$CSIMEMP = CSEMP - CSEFEMP$$

donde,

$$CSEFEMP = \sum_i PY030G_i - CSEFCP$$

representa el cálculo en la ECV de las *Cotizaciones sociales efectivas a cargo de los empleadores (recursos)* ($D_{121_r_Cne}$), siendo $PY030G_i$ las *Cotizaciones sociales efectivas del miembro i del hogar a cargo del empleador* y $CSEFCP$ las *Cotizaciones sociales efectivas de los trabajadores por cuenta propia*⁵⁸ para el total de hogares;

⁵⁸ Como ya vimos en el capítulo 1, del análisis de los microdatos de la ECV se han localizado un subconjunto de registros cuya situación profesional es la de *Empleadores o Empresarios sin asalariados* o *Trabajadores independientes* que tienen uno o varios trabajos, no declaran trabajar por cuenta ajena y tienen rentas salariales monetarias, los cuales tratamos en este trabajo como trabajadores por cuenta propia, detrayendo por tanto de dichas rentas salariales la parte correspondiente tanto en concepto de Sueldos y Salarios en efectivo ($RMCP_B$) como Cotizaciones ($CSEFCP$).

$$CSEMP = \frac{D_{12_r_Cne}}{D_{121_r_Cne}} * (CSEFEMP)$$

siendo *CSEMP* la estimación en la ECV de la partida contable *D_12_r_Cne* Cotizaciones sociales efectivas a cargo de los empleadores (recursos).

Por su parte, del total de *Rentas de la Propiedad* que el SEC 2010 tiene reguladas para el sector hogares, además de las *Rentas de la tierra* estimadas anteriormente, la ECV tampoco recoge los *Intereses pagados* (rúbrica *D_41_e_Cne* en CNE) ni las *Otras rentas de inversión* (*D_44_r_Cne*). Para su estimación en la ECV haremos uso de las variables *INTP_B* y *ORI_B*, respectivamente, teniendo que:

$$INTP_B = \frac{D_{41_e_Cne}}{(D_{4142_re_Cne} + D_{41_e_Cne})} * (IMCAP_B)$$

donde *D_4142_re_Cne* designa el saldo contable agregado de las partidas *Intereses* (*D_41_re_Cne*) y *Rentas distribuidas de las sociedades* (*D_42_re_Cne*) e *IMCAP_B* su equivalente en la ECV (*Intereses, dividendos y ganancias brutos de inversiones de capital en empresas no constituidas en sociedad*, recogidos por medio de la variable *HY090G*); y

$$ORI_B = ORI_B_EFF * (i_D) .$$

siendo,

i_D el tipo de interés de nuevas operaciones, publicado por el Banco de España, aplicado por las entidades de crédito y establecimientos financieros de crédito a los depósitos a más de 2 años de hogares e ISFLSH y sociedades no financieras residentes en la UEM; y

$$ORI_B_EFF = p5_20b_eff + p5_20c_eff + p5_23_eff + \sum_{i=1}^6 (p5_17a_i_eff + p5_17b_i_eff)$$

donde,

p5_17a_i_eff: Importe de la prima anual pagada por el miembro *i* del hogar que tiene seguros de vida suscritos por decisión propia, (*i* = 1,2,...,6);

p5_17b_i_eff: Importe del pago único de los seguros de vida suscritos por decisión propia a cargo del miembro *i* del hogar, (*i* = 1,2,...,6);

p5_20b_eff: Importe de la prima anual pagada por los hogares que tienen seguros de vida no suscritos por decisión propia;

p5_20c_eff: Importe del pago único de los seguros de vida no suscritos por decisión propia a cargo de los hogares;

p5_23_eff: Importe de los pagos que en promedio anual realizan los hogares por la contratación de otras formas de aseguramiento distintas de los seguros de vida (automóvil, decesos, accidentes, etc).

Las siguientes partidas que vamos estimar en la ECV corresponden a la Cuenta de distribución secundaria de la renta. En primer lugar, hablaremos de las Cotizaciones sociales efectivas por los hogares, *CSEFHOG*, las cuales, como ya comentamos en el capítulo II, necesitamos estimar porque en CNE se

recogen por separado mientras que en la encuesta lo hacen junto a los Impuestos sobre la renta a través de la variable *HY140G*. Para ello haremos:

$$CSEFHOG = \frac{(D_{61_er_Cne} - D_{121_r_Cne} - D_{122_r_Cne} + D_{61_r_Cne})}{D_{121_r_Cne}} * (CSEFEMP)$$

donde,

D_{61_er_Cne} son los empleos netos (empleos menos recursos) de la partida Cotizaciones Sociales Netas de CNE;

D_{121_r_Cne} son las Cotizaciones sociales efectivas de CNE;

D_{122_r_Cne} son las Cotizaciones sociales imputadas de CNE;

y *D_{61_r_Cne}* son los recursos de las Cotizaciones Sociales Netas de CNE.

La siguiente partida a estimar en la ECV para construir la Cuenta de Distribución Secundaria de la Renta, las Cotizaciones sociales imputadas a cargo de los empleadores, *CSIMEMP*, ya la estimamos dentro del bloque de variables a estimar en la Cuenta de Asignación de Renta Primaria.

Por su parte, las *Cotizaciones sociales suplementarias a cargo de los hogares (empleos)* de CNE, *D.613_e_Cne*, no se publican en CNE ni tampoco se recogen en la ECV ni en información auxiliar fusionada de la EPF y de la EFF, de modo que no tenemos forma de estimarlas.

Por último, dentro del bloque de Cotizaciones Sociales Netas de CNE, los recursos, debidos a los Servicios de los sistemas de seguros sociales, *D_{61SC_r_Cne}*, se estimarán en la encuesta por medio de la variable *OCS*, *Otras de cotizaciones sociales*, aplicando a las Cotizaciones sociales efectivas de la ECV el peso de los recursos de las Cotizaciones sociales netas de CNE sobre las Cotizaciones sociales efectivas de CNE, tal que:

$$OCS = \frac{D_{61_r_Cne}}{D_{121_r_Cne}} * (CSEFEMP)$$

Dentro del conjunto de *Prestaciones sociales recibidas* de CNE, las identificadas en la ECV como percibidas por los menores de 16 años, Prestaciones de supervivencia, Prestaciones por enfermedad e invalidez y Ayudas para estudios (*ORNC2_B*, *ORNC3_B* y *ORNC4_B*, respectivamente), no ha sido posible estimarlas ya que ninguna de las fuentes que estudiamos en este trabajo contienen información al respecto.

La ECV tampoco recoge las *Prestaciones sociales pagadas por los hogares* en CNE, debidas a las Prestaciones de otros sistemas de seguros sociales (*D.622_e_Cne*), y para estimarlas utilizaremos en la encuesta la variable *POSSS_e*, de modo que por medio del peso de dicha partida en CNE sobre el total de recursos de las Cotizaciones sociales netas en CNE, tendremos que:

$$POSSS_e = \frac{D_{62_e_Cne}}{D_{61_r_Cne}} * (OCS)$$

donde *OCS* es el equivalente en la ECV a la partida *D_{61_r_Cne}*.

A continuación, entraríamos ya en el el bloque de las *Otras transferencias corrientes*, en primer lugar, nos centraremos en las recibidas por los hogares. En este punto, la siguiente partida a estimar en la ECV son las *Indemnizaciones de seguro no vida* recibidas por los hogares como consecuencia de haber sufrido daños personales. En la ECV se recogen dentro del bloque de *Prestaciones sociales recibidas* (o bien por enfermedad o por invalidez), pero desconocemos su importe, y para estimarlas utilizaremos la rúbrica INDNOVIDA1_B. Por su parte, en CNE las indemnizaciones de seguro no vida por daños personales no se recogen como Prestaciones sociales recibidas sino como *Otras transferencias recibidas* junto a las indemnizaciones de seguros no vida recibidas por los hogares como consecuencia de los daños sufridos en sus bienes, las cuales tampoco se recogen en la ECV y que estimaremos mediante la variable INDNOVIDA2_B.

Así, la variable INDNOVIDA1_B la estimaremos con ayuda de la EFF. Así:

$$INDNOVIDA1_B = (p6_52_eff - p6_51b_eff) * 12$$

donde,

p6_52_eff: Indemnizaciones recibidas mensualmente por accidentes, salud, médicos, etc. por el hogar en el año posterior al de referencia de la encuesta.

p6_51b_eff: Aumento que se ha producido en el año posterior al de referencia de la encuesta en las Indemnizaciones recibidas mensualmente por accidentes, salud, médicos, etc. por el hogar .

Por su parte, INDNOVIDA2_B no podremos estimarla al no encontrar información al respecto en las fuentes de información estudiadas.

Otras de las partidas que necesitamos estimar en la ECV dentro del bloque de las *Otras transferencias corrientes recibidas* son las *Otras transferencias corrientes percibidas por los hogares* (OTRDOH_B), las cuales estimaremos con los datos de la EFF de los ingresos no periódicos recibidos por los hogares por lotería y juegos de azar. De esta forma, OTRDOH_B junto a las *Transferencias periódicas monetarias brutas percibidas de otros hogares* (HY080G), que sí se recogen en la encuesta, son equivalentes a los recursos de las *Transferencias corrientes diversas* de CNE (D_75_r_Cne). Así, para estimar OTRDOH_B⁵⁹ tendremos:

$$\begin{aligned} OTRDOH_B = & [(P6_60_eff - P6_59b_eff + P6_60c1_eff - P6_59bc1_eff) * 12] + \\ & + \{P6_60f_eff * [I(p6_60es1_eff = 1) + I(p6_60es2_eff = 1)]\} + \\ & + \{p6_60f_eff * [I(p6_60es3_eff = 1) + I(p6_60es4_eff = 1)]\} \end{aligned}$$

donde,

P6_60_eff: Ayuda monetaria que recibe el hogar por parte de familiares fuera del hogar o de amigos en el año posterior al de referencia de la encuesta.

⁵⁹ Se han incluido partidas equivalentes a las *Transferencias periódicas monetarias brutas percibidas de otros hogares* (HY080G) porque esta variable está muy deficientemente recogida en la ECV.

P6_59b_eff: Aumento que se ha producido en el año posterior al de referencia de la encuesta en la Ayuda monetaria que recibe el hogar por parte de familiares fuera del hogar o de amigos.

P6_60c1_eff: Ayuda monetaria que recibe el hogar por parte de una expareja con la que ya no convive en el año posterior al de referencia de la encuesta.

P6_59bc1_eff: Aumento que se ha producido en el año posterior al de referencia de la encuesta en la Ayuda monetaria que recibe el hogar por parte de una expareja con la que ya no convive.

P6_60f_eff: Incluye entre otros los ingresos por premios, loterías, etc. si $p6_60es1_eff = 1$ ó $p6_60es2_eff = 1$ ó $p6_60es3_eff = 1$ ó $p6_60es4_eff = 1$.

Respecto a las *Otras transferencias corrientes pagadas* de CNE necesitaremos estimar en la ECV las Primas netas de seguro no vida, y lo haremos por medio de la rúbrica PRIMNOVIDA_B, y los empleos de las Otras transferencias diversas debidas a las Otras transferencias no periódicas monetarias pagadas por los hogares (OTPAOH_B).

Las Primas netas de seguro no vida en la ECV, *PRIMNOVIDA_B*, las estimaremos con ayuda de la variable *p5_23_eff* de la EFF como:

$$PRIMNOVIDA_B = \frac{p5_23_eff}{\sum_{EFF} p5_23_eff} * PRIMNOVIDA_Cne$$

donde,

p5_23_eff: Importe de la prima anual de otras formas de aseguramiento (pólizas sanitarias, hogar, automóvil, etc) del hogar.

PRIMNOVIDA_Cne: Primas netas de seguro no vida de la CNE.

Por su parte, las Otras transferencias no periódicas monetarias pagadas por los hogares, *OTPAOH_B*, las estimaremos como:

$$OTPAOH_B = \left[\left(\frac{D_75_e_Cne}{D_75_r_Cne} \right) * (TPRDOH_B + OTRDOH_B) \right] - TPPAOH_B$$

donde,

D_75_e_Cne: son los empleos de la partida *Transferencias corrientes diversas* de CNE.

D_75_r_Cne: son los recursos de la partida *Transferencias corrientes diversas* de CNE.

(TPRDOH_B + OTRDOH_B): es el equivalente en la ECV a los recursos de las *Transferencias corrientes diversas* de CNE.

TPPAOH_B: es la parte de los empleos de la partida *Transferencias corrientes diversas* de CNE recogida en la ECV a través de la variable HY130G, *Transferencias periódicas monetarias brutas abonadas a otros hogares*.

Después de llevar a cabo la estimación de las partidas no disponibles en la ECV haciendo uso de las fuentes de información fusionadas, nos planteamos acometer otro de los objetivos planteados fruto de aquellas, el diseño y construcción de un conjunto de Indicadores de Incongruencia, los cuales nos

sevirán para corregir la ocultación de ingresos en los hogares de la ECV detectada en el Capítulo II de este trabajo.

5. Evidencias directas y circunstanciales acerca del fenómeno de la subdeclaración: Indicadores de incongruencia

Una vez que disponemos en la ECV de todas las partidas equivalentes a las que publica la CNE como partes integrantes de la Renta Disponible de los Hogares, en este apartado analizaremos cómo a partir de la información fusionada a los hogares de la ECV Base 2013 en el apartado 3 de este capítulo -procedente de la EPF y de la EFF-, y de las relaciones que se desprenden de la Teoría Económica entre la renta, el consumo y la riqueza -estudiadas en el apartado 2.1 del mismo- podemos tratar de construir un conjunto de indicadores de incongruencia que nos ayudarán a aflorar parte de la subdeclaración contenida en la ECV y, en consecuencia, con ayuda de ellos, a estimar una distribución personal de la renta en España más cercana a la realidad en dicha encuesta. La idea es que si entre una determinada variable de ingresos y otra variable informativa de algún otro aspecto económico del hogar se espera una cierta relación, positiva o negativa (por alguno de los motivos estudiados en el apartado 2.1), cuanto más contraria a ella sea la conducta observada de un hogar (en base a la información declarada), mayor fuerza tomará la sospecha de que dicho hogar está presumiblemente declarando mal sus ingresos. Y cuanto más relaciones contradictorias encontremos, mayor será la confianza sobre lo sospechado. Pretendemos, en suma, construir indicadores que permitan medir objetivamente lo que nos manifiestan ciertos signos externos del hogar que deberían estar relacionados con sus ingresos.

A continuación, en este apartado, iniciamos el camino para la búsqueda de estos indicadores de incongruencia, por tipo de renta, empezando por los componentes de la Cuenta de Asignación de la Renta Primaria y pasando por los de la Cuenta de Distribución Secundaria de la Renta hasta acabar en el saldo de ésta, la Renta Disponible. Y para ello, recogemos aquí las relaciones sobre las que hemos indagado y construido dichos indicadores en el contexto de este trabajo, como veremos en el capítulo siguiente.

Comenzamos con los componentes del Excedente Bruto de Explotación derivado de los servicios de alquiler de las viviendas ocupadas por los hogares que son sus propietarios:

- Para contrastar la Producción relacionada con el Alquiler imputado a la vivienda principal del hogar (AIVPP_P) cuando la ocupa en propiedad (regteneviv=1) proporcionada por la ECV, podemos pensar que ésta mantendrá una relación de proporcionalidad directa⁶⁰ con la variable

⁶⁰ En todo este apartado, para designar las relaciones de proporcionalidad directa entre dos variables utilizamos el signo (+) entre ellas; y para las inversas, intercalaremos el signo (-) entre ellas. Asimismo, para cada indicador de incongruencia que queramos construir con base en alguna relación, (por ejemplo, un

de la EFF *Valor actual de la vivienda principal* (P2_5_eff) cuando *la posee en propiedad* (P2_1_eff=2). Así, el indicador de incongruencia calculado para comprobar la adecuación de lo declarado como AIVPP_P lo denotaremos por B2_001 y se puede obtener mediante la siguiente relación:

$$[AIVPP_P * I(\text{regteneviv} = 1)](+) VC_B2_001$$

siendo

$$VC_B2_001 = P2_5_eff * I(P2_1_eff = 2)$$

- En lo que respecta a los Consumos Intermedios asociados a dicho Alquiler imputado (AIVPP_CI) de la ECV:
 - o Los Intereses pagados de préstamos hipotecarios (AIVPP_CI_IPPH) presentes cuando el hogar ocupa la vivienda habitual en propiedad pero con hipoteca (regteneviv=2), podemos pensar que mantendrán una relación de proporcionalidad directa con la variable de la EFF equivalente⁶¹ (IPPH_B_eff) cuando *la posee en propiedad* (P2_1_eff=2). Así, el indicador de incongruencia calculado para comprobar la adecuación de AIVPP_P lo denotaremos por B2_002 y se puede obtener mediante la siguiente relación:

$$[AIVPP_CI_IPPH * I(\text{regteneviv} = 2)](+) VC_B2_002$$

siendo

$$VC_B2_002 = IPPH_B_eff * I(P2_1_eff = 2)$$

- o Respecto a los Consumos Intermedios de los servicios de Alquiler imputado atribuidos a los Servicios de Seguros y Planes de Pensiones (AIVPP_CI_Seg) presentes cuando el hogar ocupa la vivienda habitual en propiedad pero con hipoteca (regteneviv=2), podemos pensar que éste mantendrá una relación de proporcionalidad directa con la variable de la EFF *Importe prima anual de Seguros de Vida No voluntarios* (P5_20b_eff) o, por su parte, con *Importe pago único de Seguros de Vida No voluntarios* (P5_20c_eff), siempre que el hogar posea la vivienda en propiedad (P2_1_eff=2). El indicador de incongruencia así calculado para comprobar la adecuación de AIVPP_P lo denotaremos por B2_013 y se puede obtener mediante la siguiente relación:

$$[AIVPP_CI_Seg * I(\text{regteneviv} = 2)](+) VC_B2_013$$

indicador VV_XXX que queramos construir para una partida que presenta alguna relación con una cierta variable de control que empleamos para calcularlo), notaremos a dicha variable de control (VC) con el mismo nombre que el indicador, pero anteponiéndole el prefijo "VC" (para el ejemplo, VC_VV_XXX)

⁶¹ Parar ver en detalle el cálculo de las variables de CNE en la EFF, de forma análoga al que se ha desarrollado en la ECV en el Capítulo 2 de este trabajo, se remite al programa informático desarrollado a tal efecto, que se adjunta en el Anexo 1.c. de este trabajo.

siendo

$$VC_B2_013 = (P5_20b_eff + P5_20c_eff) * I(P2_1_eff = 2)$$

- o Por otra parte, la lógica nos dice que también podemos contrastar la partida AIVPP_CI_Seg con el importe de los intereses pagados de préstamos hipotecarios de la EFF, para lo que utilizaremos el indicador de incongruencia B2_003, obtenido de forma que:

$$[AIVPP_CI_Seg * I(regteneviv = 2)](+) VC_B2_003$$

siendo

$$VC_B2_003 = IPPH_B_eff * I(P2_1_eff = 2)$$

- o Por último, cuando tenemos más de un indicador de incongruencia para una variable, como en este caso para la variable AIVPP_CI_Seg, podemos extraer otro indicador sintético de estos mediante la primera componente principal (CP1), en este caso de B2_013 y B2_003. Lo llamaremos indicador B2_000 y lo expresaremos como:

$$B2_000 = CP1(B2_003 ; B2_013);$$

- Los Otros impuestos netos sobre la producción asociados al Alquiler imputado relativo a la vivienda principal del hogar (AIVPP_P) cuando la ocupa en propiedad (regteneviv=1), estimados en la ECV por medio de la variable AIVPP_OIN, podemos contrastarlos suponiendo que mantendrán una relación de proporcionalidad directa con la variable de la EFF *Valor actual de la vivienda principal* (P2_5_eff) cuando la posee en propiedad (P2_1_eff=2). Así, el indicador de incongruencia calculado para comprobar la adecuación de AIVPP_OIN lo denotaremos por B2_004 y se puede obtener mediante la siguiente relación:

$$[AIVPP_OIN * I(regteneviv = 1)](+) VC_B2_004$$

siendo

$$VC_B2_004 = P2_5_eff * I(P2_1_eff = 2)$$

- Por último, el valor del Consumo de Capital Fijo atribuible al Excedente de Explotación del Alquiler imputado relativo a la vivienda principal del hogar (AIVPP_P) en propiedad (regteneviv=1), estimado en la ECV por medio de la variable CCFEE, podemos controlarlo a través de la suposición de que guardará una relación de proporcionalidad directa con el valor de la misma variable estimado en la EFF. Para ello calcularemos en la EFF dicho CCFFE -bajo la rúbrica B2_005- siguiendo la lógica de los contadores nacionales, es decir, calculando la depreciación a partir del valor corriente del activo. En CNE se valora el activo a partir del precio de compra de la propiedad (P2_4_eff) en el

año en que la obtuvieron ($P2_3_eff$) por la tasa acumulada de inflación. Suponiendo un periodo de amortización de la propiedad de 50 años y un incremento del Índice de Precios al Consumo (IPC) medio anual del grupo de Vivienda del +3,29% (en los últimos 15 años se incrementó un 62,4% según el INE, lo que supone un aumento medio del 3,29% anual). De forma que:

$$[CCFEE * I(\text{regteneviv} = 1)](+) VC_B2_005$$

siendo

$$VC_B2_005 = P2_4_eff * (1 + \overline{IPCAnual_{vivienda}})^{**} (ano - P2_3_eff) * \min[1; (ano - P2_3_eff)/50] * I(P2_1_eff = 2)$$

Una vez presentados los indicadores de incongruencia para el EBE continuamos con los relativos a la Renta Mixta:

- Para contrastar el valor de la Renta Mixta Neta de la ECV (B_3n_Bdh) asimilable a la partida B_3n que publica CNE podemos elaborar tres indicadores de incongruencia:
 - o En primer lugar, el que resulta de estimar esta misma partida en la EFF ($B3_001$) a partir de los Ingresos brutos por cuenta propia de la EFF ($RMIX_B_eff$) y los Ingresos por Alquiler de Propiedades, excluido terrenos ($IMALQPRO_eff$), en cuyo caso mantendrán una relación de proporcionalidad directa:

$$B_3n_Bdh (+) VC_B3_001$$

donde

$$VC_B3_001 = RMIX_B_eff + IMALQPRO_eff$$

- o Por otra parte, también podemos contrastar B_3n_Bdh por medio del indicador $B3_002$ diseñado como agregación de las variables de la EFF *Valor de cuentas y depósitos utilizables para realizar pagos* ($V_CUEYDEP_eff$) y del *Importe inicial Préstamos para hogares o empresas individuales no destinados a vivienda* ($P3_5_i_eff$, $i=1,...,8$), como indicador de solvencia, y también por medio de una relación de proporcionalidad directa:

$$B_3n_Bdh (+) VC_B3_002$$

donde

$$VC_B3_002 = V_CUEYDEP_eff + \sum_{i=1}^8 P3_5_i_eff$$

- o Por último, cuando consecuencia de tener más de un indicador de incongruencia para una esta variable, podemos extraer otro

indicador sintético de estos, el indicador B3_000, a través de la primera componente principal de B3_001 y B3_002. De forma que:

$$B3_{000} = CP1(B3_{001} ; B3_{002})$$

- En relación al Consumo de Capital Fijo de la Renta Mixta estimado en la ECV (CCFRM) el contraste se hará construyendo un indicador de incongruencia (B3_003) de forma similar al diseñado para contrastar el CCFFE, comentado anteriormente. En este caso la depreciación se calculará sobre el valor inicial de los negocios ($P4_{108_i_eff}$, $i=1,...,6$) teniendo en cuenta su año de comienzo ($P4_{107_i_eff}$, $i=1,...,6$), actualizado con el IPC general y considerando un periodo de amortización de los activos fijos del empresario autónomo de 20 años. Es decir:

$$CCFRM(+) VC_{B3_{003}}$$

siendo

$VC_{B3_{003}}$

$$= \sum_{i=1}^6 \left\{ [P4_{108_i_eff} * (1 + \overline{IPC_{Anual}}) * (ano - P4_{107_i_eff})] * \right. \\ \left. * \min[1; (ano - P4_{107_i_eff})/20] \right\}$$

En todos estos casos relacionados con la corrección de la Renta Mixta, el filtro considerado para la aplicación de dicha corrección es que $B_{3n_Bdh} \neq 0$ o bien que $RMIX_B_eff + IMALQPRO_eff \neq 0$.

Dentro de la partida de Remuneración de Asalariados:

- Para contrastar el componente Sueldos y salarios calculado en la ECV ($D_{11_r_Bdh}$) según las normas de CNE construiremos el indicador de incongruencia D1_001 a partir de la estimación de la partida Sueldos y Salarios en la EFF mediante la agregación de la Renta Monetaria y No Monetaria por Cuenta Ajena ($RMCA_eff$ y $RNMCA_eff$), suponiendo que mantiene con $D_{11_r_Bdh}$ una relación de proporcionalidad directa:

$$D_{11_r_Bdh} (+) VC_{D1_{001}}$$

donde

$$VC_{D1_{001}} = RMCA_eff + RNMCA_eff$$

- Por su parte, para contrastar el componente Cotizaciones sociales efectivas a cargo de los empleadores calculadas en la ECV de forma equivalente a CNE por medio de la variable $D_{121_r_Bdh}$, construiremos el indicador de incongruencia D1_002 a partir de la variable Sueldos y Salarios de la ECV estimada también de acuerdo a CNE ($D_{11_r_Bdh}$), y supondremos que entre ellos la relación de proporcionalidad que guardan es directa:

$$D_{121_r_Bdh} (+) VC_{D1_{002}}$$

donde

$$VC_D1_002 = D_11_r_Bdh$$

- Finalmente, el último de los componentes de la Remuneración de Asalariados que se va a contrastar son las *Cotizaciones sociales imputadas a cargo de los empleadores*, calculadas en la ECV según el marco de CNE por medio de la variable $D_122_r_Bdh$. Para ello, construiremos el indicador de incongruencia $D1_003$ a partir de la variable Sueldos y Salarios calculada en la ECV ($D_11_r_Bdh$), pero considerada per capita en el hogar. En este caso concreto supondremos que la relación de proporcionalidad que existe entre ellos es inversa, de forma que:

$$D_122_r_Bdh (-) VC_D1_003$$

donde

$$VC_D1_003 = D_11_r_Bdh /nmh$$

En todos estos casos relacionados con la corrección de la Remuneración de Asalariados, el filtro considerado para la aplicación de dicha corrección será: $D_11_r_Bdh > 0$, o bien que $(RMCA_B_eff + RNMCA_B_eff > 0)$.

Posteriormente entramos en el bloque de las Rentas de la Propiedad, donde hemos diseñado los siguientes indicadores de incongruencia para los siguientes componentes:

- Para la partida que recoge conjuntamente los Intereses recibidos por los hogares, así como las Rentas distribuidas de las sociedades estimadas en la ECV bajo la rúbrica $D_4142_r_Bdh$, su valor se contrastará a través de los siguientes indicadores de incongruencia:
 - o $D4_001$, construido a partir de la estimación de esta misma partida en la EFF ($IMCAP_B_eff$), a través de una relación directa:

$$D_4142_r_Bdh (+) VC_D4_001$$

con $VC_D4_001 = IMCAP_B_eff$;

- o $D4_002$, construido a partir de la variable valor total de los activos financieros del hogar ($actfinanc_eff$) proporcionada por la EFF, manteniendo también una relación directa:

$$D_4142_r_Bdh (+) VC_D4_002$$

con $VC_D4_002 = actfinanc_eff$;

- o y $D4_000$, construido como primera componente principal de $D4_001$ y $D4_002$, tal que:

$$D4_000 = CP1(D4_001 ; D4_002).$$

En cada uno de los casos anteriores, la corrección de la partida $D_4142_r_Bdh$ estará sujeta a $D_4142_r_Bdh \neq 0$ ó $IMCAP_B_eff \neq 0$ ó bien $actfinanc_eff \neq 0$.

- Por su parte, el valor de los Intereses pagados por los hogares estimados en la ECV ($D_{41_e_Bdh}$), lo contrastaremos por medio del indicador de incongruencia $D4_003$ que recoge la relación directa de esta partida con la variable de la EFF denominada *valor de deudas no inmobiliarias* ($deuotros_eff$):

$$D_{41_e_Bdh} (+) VC_{D4_003}$$

siendo $VC_{D4_003} = deuotros_eff$ y su aplicación sujeta a $D_{41_e_Bdh} \neq 0$ ó $deuotros_eff \neq 0$.

- Para la validación de la partida Otras Rentas de Inversión estimada en la ECV por medio de $D_{44_r_Bdh}$ construiremos el indicador de incongruencia $D4_004$ que recoge la relación directa de $D_{44_r_Bdh}$ con la suma de las variables de la EFF *valor de fondos de inversión* ($allf_eff$), *valor de seguros de vida* ($valor_eff$) y *valor de planes de pensiones* ($valor_eff$), de forma que:

$$D_{44_r_Bdh} (+) VC_{D4_004}$$

siendo $VC_{D4_004} = allf_eff + valseg_eff + valor_eff$ y su aplicación sujeta a $D_{44_r_Bdh} \neq 0$ ó $actfinanc_eff \neq 0$.

- Finalmente, la última partida perteneciente a las Rentas de la propiedad cuyo valor vamos a contrastar corresponde al saldo de la partida Rentas de la tierra ($D_{45_re_Bdh}$). Para ello, haremos uso del indicador $D4_006$, construido a partir de la estimación de la partida en cuestión en la EFF (Rentas del hogar relacionadas con el alquiler de terrenos $IMALQTER_B_eff$), a través de una relación directa:

$$D_{45_re_Bdh} (+) VC_{D4_006}$$

con $VC_{D4_006} = IMALQTER_B_eff$ y bajo la restricción $D_{45_re_Bdh} \neq 0$ ó $IMALQTER_B_eff \neq 0$.

A continuación, pasamos a describir la serie de indicadores de incongruencia que nos servirán para corregir algunas de las partidas de la Cuenta de Distribución Secundaria de la Renta estimadas por en la ECV. El primero de ellos corresponde a los Impuestos sobre la renta ($D_{51_e_Bdh}$) y contrastará el resultado de éstos por medio del Saldo de rentas primarias netas de la propia encuesta (B_{5n_Bdh}). Lo hemos denominado $D5_001$ y la relación de dicho saldo con $D_{51_e_Bdh}$ será directa, tal que:

$$D_{51_e_Bdh} (+) VC_{D5_001}$$

con $VC_{D5_001} = B_{5n_Bdh}$.

Para la siguiente partida, los Otros impuestos corrientes ($D_{59_e_Bdh}$), proponemos para su corrección el indicador de incongruencia $D5_002$, fundamentado en la relación directa de la variable de la EFF *riquezabr_eff*, *valor de la riqueza bruta del hogar* (total de activos reales y financieros):

$$D_{59_e_Bdh} (+) VC_{D5_002}$$

siendo $VC_{D5_002} = riquezaabr_eff$.

Por su parte, los empleos de la partida Cotizaciones sociales netas ($D_{61_e_Bdh}$) vamos a corregirlos en la ECV por medio de los Sueldos y Salarios estimados en la propia encuesta ($D_{11_r_Bdh}$). Para tal fin, hemos diseñado el indicador de incongruencia $D6_001$, que construimos a partir de la que suponemos una relación directa de esta partida con $D_{61_e_Bdh}$, tal que:

$$D_{61_e_Bdh} (+) VC_{D6_001}$$

con $VC_{D6_001} = D_{11_r_Bdh}$, que aplicaremos bajo la restricción $D_{61_e_Bdh} > 0$ ó $D_{11_r_Bdh} > 0$, siendo $D_{11_r_Bdh}$ la Remuneración de Asalariados de la ECV.

Los recursos de las Cotizaciones sociales netas estimados en la ECV ($D_{61_r_Bdh}$) supondremos que van a mantener una relación de proporcionalidad directa con la variable de la EFF *número de trabajadores de los negocios de la persona de referencia* ($N_TRABNEG_eff$). De esta forma, el indicador de incongruencia así diseñado ($D6_002$) se obtiene de la relación:

$$D_{61_r_Bdh} (+) VC_{D6_002}$$

con $VC_{D6_002} = N_TRABNEG_eff$ que aplicaremos sujeto a $D_{61_r_Bdh} > 0$ ó $N_TRABNEG_eff > 0$.

En relación a la partida correspondiente a los recursos de las Prestaciones sociales distintas de transferencias sociales en especie estimada en la ECV ($D_{62_r_Bdh}$), el contraste se hará construyendo el indicador de incongruencia $D6_003$ a partir de la estimación de la partida aproximada en la EFF ($PRESOC_B_m_eff$), mediante una relación directa. Como $PRESOC_B_m_eff$ y $D_{62_r_Bdh}$ no contienen los mismos conceptos, para homogeneizarlas restaremos a la segunda las Cotizaciones Sociales imputadas a Cargo de los empleadores ($CSIMEMP$) y las Indemnizaciones Brutas de Seguros No Vida ($INDNOVIDA1_B$) estimadas en la ECV. Así:

$$(D_{62_r_Bdh} - CSIMEMP - INDNOVIDA1_B) (+) VC_{D6_003}$$

con $VC_{D6_003} = PRESOC_B_m_eff$, que aplicaremos bajo la restricción $D_{62_r_Bdh} > 0$ ó $PRESOC_B_m_eff > 0$.

De forma análoga al procedimiento utilizado para la corrección de la partida $D_{61_r_Bdh}$, en el caso de los empleos de las Prestaciones sociales distintas de transferencias sociales en especie estimados en la ECV diseñamos el indicador de incongruencia $D6_004$ a partir de la variable $N_TRABNEG_eff$ que supondremos que van a mantener una relación de proporcionalidad directa con la partida a corregir, tal que:

$$D_{62_e_Bdh} (+) VC_{D6_004}$$

con $VC_D6_004 = N_TRABNEG_eff$, que aplicaremos sujeto a $D_62_e_Bdh > 0$ ó $N_TRABNEG_eff > 0$.

Una vez analizadas las correcciones que vamos a aplicar a las Prestaciones sociales, pasamos al bloque de las Otras transferencias corrientes. En primer lugar, la partida Indemnizaciones de seguro no vida estimada en la ECV ($D_72_r_Bdh$) vamos a contrastarla mediante el valor Primas netas de seguros no vida pagadas por los hogares también estimadas en dicha encuesta por medio de la variable $PRIMNOVIDA_B$, suponiendo que entre ellas existe una relación directa. Para ello nos serviremos del indicador de incongruencia $D7_001$:

$$D_72_r_Bdh (+) VC_D7_001$$

siendo $VC_D7_001 = PRIMNOVIDA_B$, que aplicaremos sujeto a $D_72_r_Bdh > 0$ ó $INDNOVIDA1_B > 0$.

Por su parte, para corregir los recursos de la partida Transferencias corrientes diversas estimados en la ECV ($D_75_r_Bdh$) hemos creado el Indicador de incongruencia $D7_002$ a partir de la variable de la EFF Transferencias corrientes brutas recibidas de otros hogares ($TCDOH_B_eff$) suponiendo que con aquélla existe una relación de proporcionalidad directa, de forma que:

$$D_75_r_Bdh (+) VC_D7_002$$

siendo $D7_002 = TCDOH_B_eff$, que aplicaremos sujeto a $D_75_r_Bdh > 0$ ó $TCDOH_B_eff > 0$ ó $OTRDOH_B_eff > 0$ ó ($P1_11d_eff = 1$ and $anoe = 2011$), donde

$OTRDOH_B$ son las Rentas recibidas de Otros Hogares (ECV) y $P1_11d_eff = 1$ es el valor que toma la variable $P1_11d_eff = 1$ cuando Otros miembros de fuera del hogar ayudan económicamente (en la EFF 2011).

En el caso de las Primas netas de seguro no vida ($D_71_e_Bdh$) estimadas en la ECV supondremos que van a mantener una relación directa con la variable de la EFF Pagos anuales en otras formas de aseguramiento (no vida) ($P5_23_eff$). Así, el indicador de incongruencia así diseñado ($D7_003$) se calcula a partir de la relación:

$$D_71_e_Bdh (+) VC_D7_003$$

con $VC_D7_003 = P5_23_eff$, que aplicaremos sujeto a $D_71_e_Bdh > 0$ ó $P5_23_eff > 0$.

Por último, dentro de las Otras transferencias corrientes hemos creado el indicador de incongruencia $D7_004$ destinado a la corrección de los empleos de la partida Transferencias corrientes diversas estimados en la ECV ($D_75_e_Bdh$). Para ello, hemos supuesto que la variable de la EFF Dinero que se pasa a otras personas ajenas al hogar ($P9_4_eff$) está relacionado de forma directa con $D_75_e_Bdh$, tal que:

$$D_{75_e_Bdh} (+) VC_{D7_004}$$

con $VC_{D7_004} = P9_4_eff$, que aplicaremos sujeto a $D_{75_e_Bdh} > 0$ ó $P9_4_eff > 0$.

Finalmente, el último bloque de indicadores de incongruencia tiene como objetivo aplicar las correcciones directamente sobre el saldo de la Cuenta de distribución secundaria de la renta de los hogares, es decir, la Renta disponible del hogar. En concreto, la variable sobre la que vamos a intervenir es la *Renta disponible neta per cápita del hogar* estimada en la ECV, calculada como la *Renta disponible neta del hogar* (B_{6n_Bdh}) entre el *número de miembros de cada hogar* (nmh), es decir, B_{6n_Bdh}/nmh . Para ello, hemos elaborado diferentes tipos de indicadores de incongruencia a partir de las variables fundamentales fusionadas en el proceso de *matching* como son: el *gasto total del hogar*, tanto el proporcionado por la EPF ($GASTOH_epf$) como por la EFF ($GASTOH_eff$) y la *riqueza del hogar* de la EFF en *términos brutos* ($riquezabr_eff$) y *netos* ($riquezamet_eff$), siempre consideradas per cápita. Así las relaciones sobre las que se construyen los correspondientes indicadores de incongruencia general son las siguientes:

$$(B_{6n_Bdh}/nmh) (+) (VC_{G_001})$$

$$(B_{6n_Bdh}/nmh) (+) (VC_{G_002})$$

$$(B_{6n_Bdh}/nmh) (+) (VC_{G_003})$$

$$(B_{6n_Bdh}/nmh) (+) (VC_{G_004})$$

siendo $VC_{G_001} = GASTOH_epf/nmh$, $VC_{G_002} = GASTOH_eff/nmh$, $VC_{G_003} = riquezamet_eff/nmh$ y $VC_{G_004} = riquezabr_epf/nmh$.

Por último, hemos extraído la primera componente principal de los cuatro indicadores de incongruencia general así obtenidos para sintetizar su información en uno nuevo, el indicador G_{000} , siendo por tanto:

$$G_{000} = CP1(G_{001}; G_{002}; G_{003}; G_{004}).$$

Un resumen de la serie de indicadores de incongruencia estudiados se muestra en los cuadros 4.14.a, 4.15.a y 4.16.a del capítulo siguiente.

Una vez disponemos de información de los hogares de la ECV homogénea a la que comprende el cálculo de la Renta disponible de los hogares en CNE, y de relaciones teóricas a partir de las que calcular un conjunto de indicadores de incongruencia que nos servirán para obtener y aplicar a cada uno de los componentes de dicha Renta disponible en la ECV las correcciones encaminadas a eliminar la subdeclaración aflorada en cada uno de ellos en el capítulo II de este trabajo, en el capítulo IV que se presenta a continuación expondremos la metodología final propuesta para calcular concretamente dichos estimadores de incongruencia para cada tipo de renta declarada en cada hogar, así como emplear sus resultados para llevar a cabo la estimación de una Distribución Personal de la Renta (DPR) potencialmente más cercana a la realidad que describe agregadamente la CNE, que la que en la actualidad se desprende directamente de las ECV originales.

Capítulo IV

La estimación combinada de la Distribución Personal de la Renta (DPR)

1. La estimación de una DPR congruente: antecedentes y objetivos

Como hemos venido argumentando a lo largo de este trabajo, considerar exclusivamente la información aportada por las encuestas como expresión fidedigna de la distribución del tamaño de las rentas de las personas y familias, sin efectuar ningún tipo de corrección por subdeclaración, nos llevaría a una visión distorsionada de la realidad (en mayor grado en la base 2004, en menor grado en la base 2013), por muy buenos que fuesen los métodos aplicados para su análisis; distorsionada tanto en los niveles, como en el perfil dibujado por la distribución.

En el capítulo 5 de la publicación "Distribución Personal de la Renta en España" (Pena *et al.*, 1996) encontramos los primeros intentos de corrección de la distribución personal de la renta en España, en la línea en la que esta Memoria se ubica. En concreto, en este trabajo, las correcciones se efectúan en dos etapas: primero se obtienen los niveles agregados que debería presentar la distribución según diferentes desagregaciones (ámbitos regionales, socio-profesionales y tipo de hábitat) y, una vez fijados estos niveles, se proponen diversos procedimientos para realizar un reparto aceptable de las diferencias agregadas observadas entre los individuos.

En nuestro trabajo, hemos dedicado el segundo capítulo a construir una Cuenta de Rentas del Sector Hogares (CRSH) compatible con la información recogida en las ECV y homogénea con la CNE, con el objeto de poder establecer de la forma más precisa posible dichos niveles agregados que deberían reproducir fielmente una distribución más realista del tamaño de las rentas de los hogares españoles.

Aun conociendo las críticas que en algunos aspectos suscitan los datos de CNE, la continuidad de su publicación, el carácter público y de obligado cumplimiento de su metodología, los esfuerzos continuos de mejora que en ella se vuelcan año a año, etc. nos lleva a considerarla el marco de referencia más fiable para calcular dichas discrepancias, a pesar de ser conscientes de algunas de sus imperfecciones y en la esperanza de que el método que proponemos irá produciendo estimaciones tanto más realistas cuanto más realista sea paralelamente la CNE.

Así, una vez que hemos calculado en el capítulo II estas discrepancias entre la ECV y la CNE que el método de estimación de la Distribución del tamaño de las rentas debería corregir, en este capítulo nos centraremos, fundamentalmente, en la segunda de aquellas etapas: una vez fijados dichos niveles agregados de incongruencia por partidas de la CRSH, propondremos una metodología para realizar un reparto aceptable de dichas diferencias, tratando de dar soluciones objetivas a las decisiones un tanto subjetivas que se tenían que tomar en los métodos propuestos en Pena *et al.*, 1996. Con esta finalidad, en el capítulo III hemos allegado información auxiliar proveniente de otras fuentes estadísticas que consideramos complementarias de las ECV y que nos permitirán abordar nuestros objetivos.

En aquél trabajo, una vez determinadas las diferencias entre los niveles de las partidas de la CNE y los que daban las Encuestas, -denominadas "Rentas

ocultas"- que había que repartir entre los individuos, para cada nivel de desagregación, se hicieron varios intentos de reparto, que tenían en consideración las pautas de comportamiento que habían observado del análisis de dichas diferencias por partidas. Entre otras pautas, los autores observaron que:

"... la proporción de ingresos que representan los ingresos monetarios frente al total de ingresos per cápita es bastante alta en todos los intervalos interdecílicos, y todos los años, siendo prácticamente siempre superior al 85% de los ingresos per cápita totales, siendo prácticamente insignificantes los ingresos monetarios no regulares, salvo quizá en el último intervalo interdecílico de la EBPF-90/91, presumiblemente de forma anómala.

Así mismo, se observa cómo en el tiempo va perdiendo peso la proporción de ingresos debido a la partida por Cuenta propia con respecto a la de por Cuenta ajena, que cada vez se hace más importante, lo que no parece demasiado lógico si pensamos en la proliferación de licencias fiscales, profesionales autónomos, etc. en estos años. Además, las rentas de la propiedad aumentan con el tiempo, especialmente en los intervalos interdecílicos bajos, y prácticamente nada en los altos, lo que parece un contrasentido, si pensamos que las propiedades deben ir acumulándose lógicamente en los individuos de rentas altas.

Además, los porcentajes de participación de estos ingresos de la propiedad parecen a simple vista ridículos, si pensamos que solamente en el caso de inmuebles, ya el Ministerio de Hacienda considera que para éstos es de un 2%, al margen de otro tipo de gravámenes que recaen sobre ellos como patrimonio, contribución urbana, etc., que lógicamente la renta de los mismos debe cubrir.

Todo esto hace pensar sin duda en un alto grado de ocultación en las profesiones liberales y autónomas, así como en las rentas de la propiedad."

Con base en estos razonamientos, los autores tratan de orientar el procedimiento de corrección entonces propuesto, y dicen:

"Podría pensarse en utilizar la información recogida en dichas encuestas de Presupuestos Familiares, referentes a las partidas en que presumiblemente la ocultación es mayor (rentas de capital y mixtas) para utilizarlas en un procedimiento de imputación de cantidades ocultas a las familias. Sin embargo, dado que precisamente éstas podrían ser las cantidades más falseadas, con la información disponible de las EBPF difícilmente podríamos pasar a corregir las rentas de este tipo, en base justamente a su importancia, que prácticamente desconoceríamos, al ser presumiblemente la parte más ocultada, pudiendo tal procedimiento arrojar resultados insospechados."

Esto les obliga a plantear hipótesis sobre el comportamiento de los hogares

frente al fenómeno detectado de la ocultación, y proponen:

Como una medida simple, hemos considerado, para abordar el problema de corregir nuestros datos, la hipótesis de que la ocultación debe recaer más en las clases de mayores ingresos, ya que serán éstas las que podrán acceder más fácilmente a la propiedad de bienes, y por tanto, a mayor nivel de ingresos, mayor nivel de ocultación debemos imputarle.

Por ello hemos preferido utilizar el procedimiento más simple que con la información disponible puede ofrecernos una aproximación global válida a la imputación de ingresos ocultos a las familias: realizaremos esta imputación basándonos en la hipótesis inicial que asume que las cantidades ocultas en cada clase son proporcionales a la renta repartida en la misma. Y puesto que disponemos de las cantidades ocultas en cada región (OR_i , $i=1,2,\dots,18$), en cada categoría socio-profesional (OC_j , $j=1,2,\dots,10$), y en cada clase de hábitat (OH_k , $k=1,2,3,4$), la corrección se hará respetando la pertenencia de las familias a estas clases, y teniendo en cuenta la ocultación global que en ellas se dio."

Pero en aquella propuesta, de una u otra forma el peso de la corrección seguía recayendo sobre las declaraciones realizadas, debilitando, consecuentemente, el planteamiento.

Es éste el cambio fundamental que hemos tratado de dar en aquéllos modelos que tomamos de punto de partida, pero que pretendemos liberar de esta dependencia, empleando la información adicional allegada al proceso mediante las técnicas de fusión estadísticas planteadas en el capítulo III de esta tesis para medir el grado de incongruencia de cada dato declarado por los hogares, con respecto de ciertas variables relacionadas con el tipo de renta que se declara.

1.1. El modelo pionero de Pena *et al.*

El modelo base del que partimos fue el que finalmente adoptaron estos autores (Pena *et al.*, 1996), tras probar otras propuestas iniciales. Lo denominaron *Modelo de Corrección con Tasa de Ocultación progresiva (lineal)*.

En primer lugar, dicho modelo exigía a las correcciones que buscaban que respetaran los totales ocultos por clase (y, consecuentemente, el total oculto nacional), así como también que respetaran el principio de proporcionalidad de la corrección a los ingresos declarados (a más ingresos, más ocultación); pero, además, en esta propuesta, si notamos por O_x la cantidad oculta por un receptor de una renta x , y llamamos tasa de ocultación con respecto de su renta, $TO(x)$, al cociente:

$$TO(x) = \frac{O_x}{x}$$

Pena *et al.* exigen también que dicha tasa de ocultación crezca linealmente con las rentas declaradas. De esta forma, el modelo parte de esta hipótesis de progresividad lineal en la propensión a la ocultación a medida que los ingresos aumentan, de modo que, cuando éstos aumentan, no sólo aumente la cantidad ocultada, sino también la intensidad o proporción de la cantidad ocultada. Es decir, que la ocultación crecería con la renta (declarada) de forma cuadrática.

Bajo este supuesto, el modelo plantea encontrar unos coeficientes, a y m , que caractericen la progresividad lineal de la tasa de ocultación en función de la renta, tales que:

$$\text{la tasa de ocultación: } TO(x) = \frac{O_x}{x} = a + m \cdot x$$

$$\text{la ocultación total en la clase: } O = \sum_{clase} O_x$$

$$\text{y la cantidad ocultada por un individuo de renta } x: O_x = (a + m \cdot x) \cdot x$$

De este planteamiento se deriva que:

$$O = \sum_{clase} O_x = \sum_{clase} (a + m \cdot x) \cdot x = a \sum_{clase} x + m \sum_{clase} x^2 \Rightarrow a = \frac{O - m \sum_{clase} x^2}{\sum_{clase} x}$$

$$\text{de donde: } O_x = (a + m \cdot x) \cdot x = \frac{O - m \sum_{clase} x^2}{\sum_{clase} x} \cdot x + m \cdot x^2$$

$$\text{y la cantidad corregida: } X = x \left(1 + \frac{O - m \sum_{clase} x^2}{\sum_{clase} x} + m \cdot x \right)$$

Este modelo presenta dos casos extremos:

1º) Caso $m=0$, ó caso en el que la tasa de ocultación es constante.

En este caso, la cantidad corregida sería directamente proporcional a la declarada:

$$\text{cantidad corregida: } X = x \left(1 + \frac{O}{\sum_{clase} x} \right)$$

caso éste que los propios autores califican de “algo ingenuo”, en base a considerar que el propio Ministerio de Economía y Hacienda (en la actualidad, Ministerio de Economía, Industria y Competitividad) prevé una escala de gravamen de tipo progresivo para las rentas de las personas físicas; lo que podría justificar también una cierta progresividad en los intentos de ocultación.

2º) Caso $a=0$, ó de tasa de ocultación con máxima pendiente.

En este segundo caso, como:

$$O = \sum_{clase} O_x = \sum_{clase} (a + m \cdot x) x = m \sum_{clase} x^2$$

$$\text{la pendiente maxima seria: } m = \frac{O}{\sum_{clase} x^2}$$

$$\text{y la cantidad corregida: } X = x \left(1 + \frac{O \cdot x}{\sum_{clase} x^2} \right)$$

Los autores piensan que esta corrección grava quizás de forma insignificante a las rentas bajas, donde obviamente también existe interés por ocultar, asignando una ocultación mínima a las menores rentas declaradas (ocultación cero a la renta declarada cero).

Entre estas dos situaciones límite, la elección de una pendiente m plausible queda indeterminada. Los autores, en el trabajo pionero del nuestro, adoptan como hipótesis final de trabajo una situación intermedia entre estas dos situaciones extremas anteriores, prefijando como pendiente de la recta que modela la tasa de ocultación, una fracción $1/k$ de la máxima posible.

$$m = \frac{1}{k} \frac{O}{\sum_{clase} x^2}$$

$$\text{entonces: } O_x = (a + m \cdot x) x = \frac{O - m \sum_{clase} x^2}{\sum_{clase} x} x + m \cdot x^2 = \frac{O \cdot x}{k} \left(\frac{k-1}{\sum_{clase} x} - \frac{x}{\sum_{clase} x^2} \right)$$

$$\text{y la cantidad corregida: } X = x \left(1 + \frac{O}{k} \left\{ \frac{k-1}{\sum_{clase} x} - \frac{x}{\sum_{clase} x^2} \right\} \right)$$

Esta corrección asigna mayores niveles de ocultación sobre las clases bajas y medias a lo que lo hacía la anterior, aunque conserva la proporcionalidad de la tasa de ocultación en función de la renta, y respeta la cantidad total ocultada por cada clase.

En su trabajo, los autores consideran varios casos incluyendo los extremos y algunos intermedios para finalmente decantarse por un valor de "k" aparentemente aceptable, si bien la valoración comparativa de las correcciones obtenidas y la decisión final sobre "k" y la pendiente de la progresividad de la tasa de ocultación (pendiente igual a la media de las de las dos situaciones extremas), fue fundamentalmente subjetiva.

En nuestra opinión, este modelo pionero, base de este capítulo, adolece de algunos problemas, cuya solución pretendemos abordar con esta propuesta.

En primer lugar, y como ya hemos mencionado anteriormente en algunas ocasiones, las ocultaciones estimadas dependen exclusivamente de las rentas declaradas. Así, el procedimiento "se fía por igual" de las cantidades declaradas por las personas u hogares, sin dar lugar a considerar que existen personas y hogares responsables y cumplidores que declaran lo más correctamente que pueden, y que conviven en la misma muestra con otros menos responsables, o más olvidadizos, o simplemente menos cumplidores y solidarios que podrían tratar de ocultar gran parte de sus rentas por diversos motivos, entre los que estaría obviamente el temor fiscal.

En segundo lugar, el método de Pena *et al.* no ayuda a determinar de forma concreta la pendiente "m" de la tasa de ocultación, o lo que es igual, la fracción "k" de la máxima viable para la tasa de ocultación, dejando este grado de libertad a la subjetividad del investigador; lo que siempre podría ser cuestionable desde otras posiciones también subjetivas.

Por último, una forma lineal de la tasa de ocultación como función de las rentas declaradas es cuestionable, especialmente para las rentas bajas. La linealidad de la tasa de ocultación basada en la renta declarada, conduce, como podemos observar en las expresiones anteriores, a funciones parabólicas crecientes con una corrección mínima en la menor de las rentas declaradas, pudiendo ocurrir que un hogar que no declare, o que oculte todas sus rentas, no vería corregida su renta en absoluto. Esto ha sido lógicamente criticado por los autores que advierten este hecho, incluido los propios autores del trabajo pionero.

Sin embargo, el motivo de no haber abordado directamente una forma funcional de "U", más plausible, para la corrección de las rentas en función de sus rentas declaradas, no era otro que no complicar aún más el tercero de los problemas antes advertidos: si con una tasa de ocultación lineal ya dependía de la decisión subjetiva del investigador prefijar el parámetro "k" que marcaba la intensidad de la propensión a la ocultación, una forma parabólica de ésta conduciría a tener dos parámetros libres que habría que prefijar de forma subjetiva.

1.2. El modelo de Ortega y Callealta como precedente de la utilización de indicadores de rareza para el estudio de la ocultación

Un intento de abordar este problema, para tratar de determinar dicha forma funcional de "U" de forma objetiva se presenta en Ortega y Callealta,

2002. Como advertían los autores del trabajo pionero antes citado (Pena *et al.*, 1996), a causa de la poca información de partida de la que se disponía, las ocultaciones obtenidas mediante unas *Tasas Progresivas Lineales* para rentas declaradas pequeñas serían cercanas a cero (siendo cero para rentas nulas), no siendo del todo coherentes con las previsiones racionales. Parece lógico pensar que cuando las rentas son relativamente pequeñas los individuos tenderán a ocultarlas casi en absoluto o, al menos, infradeclarando considerablemente su valor. Más concretamente, refiriéndose a las rentas declaradas muy bajas, estos autores defienden que:

“...una Tasa de Ocultación debería ser no acotada para estos valores, permitiendo corregir rentas declaradas casi nulas mediante rentas de supervivencia, como mínimo. Así, una mejor aproximación de las ocultaciones con respecto a las rentas declaradas, debería seguir un modelo parabólico general en forma de “U”, de manera que la Función de Ocultación pudiera ser representada por un polinomio de segundo grado en su expresión más general. Es decir;

$$FO(x) = A + B x + C x^2$$

para valores A, B y C, de manera que la función Tasa de Ocultación sería de la forma:

$$TO(x) = (A / x) + B + C x$$

donde el nuevo modelo produciría una Tasa de Ocultación de forma hiperbólica cerca de 0 y convergente a una recta, de pendiente C, para valores grandes de x.

Pero, como advertimos anteriormente, si ya quedaba indeterminada la pendiente (o el parámetro “k”) cuando la tasa de ocultación tenía forma lineal, en este último caso, determinar la Ocultación mediante esta forma que introduce un componente adicional exige imponer alguna otra restricción o aportar al modelo información externa adicional.

Para aportar esta necesaria información extra, Ortega y Callealta proponen un procedimiento en el que parten de los resultados obtenidos en Pena *et al.* (1996) y aplican las teorías propuestas en Ortega (2000) para el estudio de la rareza de las observaciones atípicas, empleando la información así obtenida para cada caso como información extra a partir de la que ajustar la función de tipo cuadrática propuesta para la Ocultación.

Para ello, consideran las seis clases de *Ingresos Monetarios*⁶² (I_k para $k=1,2,...,6$) corregidos por el *Coeficiente de Cobertura* deducido de las

⁶² Los Ingresos Monetarios se dividen en seis tipos diferentes que son: Ingresos por Cuenta Ajena (Ica), Ingresos por Cuenta Propia (Icp), Ingresos por Rentas del Capital (Irc), Ingresos por Transferencias Regulares (Itr), Ingresos por Transferencias Ocasionales (Ito) y Otros Ingresos (Io). Además, otras notaciones del modelo son: Ingresos Monetarios (IM) y los Ingresos no Monetarios (InM), restando las Cotizaciones a la Seguridad Social (Co), los Impuestos (Im) y los Desembolsos (De), donde estos últimos se calculan como la suma de los Intereses Pagados (Ip), las Primas de Seguros (Ps), y las Transferencias ya sean Regulares a Hogares (Trh), Ocasionales a Hogares (Toh) o a Instituciones (Ti).

Encuestas Básicas de Presupuestos Familiares (EBPF) por comparación con los gastos totales de consumo de las familias, y modelizan la variable Y , suma de dichos ingresos monetarios corregidos más la ocultación detectada por el procedimiento de Pena *et al.*, como una combinación lineal de éstas, de la forma:

$$Y = \sum_{k=1}^6 I_k + ocul = \sum_{k=1}^6 \beta_k I_k (+ \text{residuo})$$

Los parámetros β_k (para $k=1,2,\dots,6$) serían indicadores relativos del nivel medio de subdeclaración imputable inicialmente a cada tipo de ingreso en la subpoblación estudiada, mientras que el *residuo* complementaría a éstos recogiendo el efecto adicional (positivo o negativo) sobre los mismos de la restante ocultación de carácter marcadamente no lineal.

Para cada clase (región), empleando las teorías robustas propuestas en Ortega (2000) y, según se define allí, para cada observación, se determina una distancia llamada *Distancia por Truncamiento* (DT_i^2) que, cuanto mayor es, nos da un mayor indicio de que dicha observación es una observación atípica para el modelo, presumiblemente, por contener una componente ocultación no lineal considerable. Por tanto, el objetivo de este primer paso del procedimiento propuesto era, exclusivamente, determinar el grado de rareza de las observaciones frente a la linealidad del modelo.

Con el propósito de estimar el modelo con una mínima influencia de las posibles observaciones atípicas y, posteriormente, proceder a la descomposición en una parte lineal y otra no lineal, Ortega y Callealta estiman unos nuevos modelos que potencian la presencia de aquellas observaciones que son consideradas como genuinas a partir de los modelos anteriores (las que presentan un grado pequeño de ocultación lineal), y tienen menos en cuenta las observaciones atípicas en aquéllos (las que presentan un mayor grado de ocultación específica no lineal). Así, plantean nuevos modelos lineales para cada clase j , donde cada observación i se pondera con el producto de su *Factor de Elevación* correspondiente (FE_i) y la raíz cuadrada del cociente $1/DT_i^2$ (ya que cuanto mayor es el valor de la *Distancia por Truncamiento*, en mayor medida se supone que la observación i -ésima es una observación atípica), de la forma:

$$Y = \sum_{k=1}^6 \beta_k I_k$$

$$\text{Ponderaciones: } w_{2i} = FE_i (1/DT_i^2)^{1/2} \quad \forall i \in N_j$$

Y , mediante Mínimos Cuadrados Ponderados, se obtienen estimaciones de Y , que denotaremos por \hat{Y} , y unos coeficientes asociados a cada tipo de ingreso considerado ($\hat{\beta}_k$ para $k=1,2,\dots,6$).

Estas estimaciones nos permiten construir la variable *ocul* como suma de dos cantidades: la primera, que denotaremos por ε' , hace referencia a la ocultación lineal asumida por el modelo (que distingue las participaciones de los distintos tipos de ingresos), y se define mediante los parámetros de regresión estimados:

$$\varepsilon' = \sum_{k=1}^6 (\hat{\beta}_k - 1) l_k$$

y la segunda, denotada por ε , representa la parte de la ocultación no lineal del mismo, y que no sería otra que el error residual de regresión propiamente dicho:

$$\varepsilon = Y - \hat{Y}$$

pudiéndose demostrar que la variable *ocul*, definida como diferencia entre las rentas corregidas y declaradas allí definidas, *rc* y *rd*, equivale a la suma de las dos variables anteriores, es decir:

$$ocul = rc - rd = (Y + ldc) - \left(\sum_{k=1}^6 l_k + ldc \right) = Y - \sum_{k=1}^6 l_k = \varepsilon' + \varepsilon$$

por lo que la *Ocultación* se redistribuye en una parte comúnmente explicada para todos los miembros de la muestra como indicador de la ocultación promedio en los diferentes tipos de ingresos (ε'), y una específica dependiente del individuo (ε).

Es decir, dichos autores descomponen la *Ocultación* propuesta por Pena et al., mediante este procedimiento, en una suma de dos cantidades: una que representa la parte lineal de la misma, y otra que representa la parte no lineal.

Será con esta parte específicamente no lineal de la ocultación (ε), con la que se ajustará el modelo parabólico buscado respecto de la *Renta Declarada (rd)*, de manera que sea posible obtener una redistribución de las ocultaciones iniciales dadas por Pena et al. y, al mismo tiempo, se otorgue mayor importancia a las observaciones atípicas en los modelos lineales. Así, para cada clase *j*, se definen los modelos:

$$\varepsilon = A_j + B_j rd + C_j rd^2$$

$$\text{Restricción: } \sum_{i \in N_j} \hat{\varepsilon}_i FE_i = OR_j - \sum_{i \in N_j} \varepsilon'_i \cdot FE_i$$

$$\text{Ponderaciones: } w_{3i} = FE_i (DT_i^2)^{1/2}, \quad \forall i \in N_j$$

donde $\hat{\varepsilon}$ representa la estimación buscada de ε , y OR_j representa la *Ocultación para la Región j-ésima* con respecto a la Contabilidad Nacional.

Como resultado, para cada clase, se estima el modelo mediante Mínimos Cuadrados Ponderados, considerando la restricción, y se obtienen los coeficientes \hat{A}_j , \hat{B}_j y \hat{C}_j mediante los cuales es posible construir una estimación de ε ($\hat{\varepsilon}$), que nos proporcione una redistribución de la *Ocultación (reocul)* calculada en Pena et al., de la forma:

$$reocul = \varepsilon' + \hat{\varepsilon}$$

Finalmente, Ortega y Callealta utilizan esta redistribución de las ocultaciones (*reocul*) para estimar la forma de la *Función de Ocultación* en cada clase, con el fin de poder predecir la *Ocultación* para un individuo de

una determinada región con una determinada *Renta Declarada*. Y lo hacen mediante el modelo restringido:

$$r = reocul = A'_j + B'_j rd + C'_j rd^2$$

$$\text{Restricción: } \sum_{i \in N_j} \hat{r}_i FE_i = OR_j$$

$$\text{Ponderaciones: } w_{4i} = FE_i, \quad \forall i \in N_j$$

donde \hat{r} representa la estimación de la variable *reocul* de manera que se respete la información proporcionada por la CNE (mediante la restricción), y donde se conserve la estructura de la EBPF (mediante las ponderaciones). Las estimaciones de Mínimos Cuadrados Ponderados resultantes para los coeficientes del modelo, considerando la restricción, permiten definir la *Función de Ocultación* para la clase *j*-ésima de la forma:

$$FO(rd) = \hat{A}'_j + \hat{B}'_j rd + \hat{C}'_j rd^2$$

Es éste, por tanto, un primer estudio en la línea que ahora proponemos en este trabajo: hacer pivotar la corrección de los datos en la rareza o atipicidad de los datos frente a la relación lineal esperada (cuando lo que se trataba era estimar la parte no lineal de la ocultación), y que en nuestro modelo haremos pivotar sobre los que llamaremos indicadores de incongruencia frente a las relaciones esperadas con otras variables que tienen relación directa ó inversa con las partidas de ingresos que han sido presuntamente mal declaradas.

1.3. El modelo de Callealta y Romero, como precedente de la utilización de fusiones estadísticas para el estudio de la ocultación

Sin embargo, todos estos estudios anteriores consideran la renta declarada como único dato para tratar de corregir las distribuciones personales de renta de personas y hogares. Y a pesar del esfuerzo de este último trabajo por extraer de los datos la máxima información al respecto, siempre puede imaginarse insuficiente, siendo muy conveniente tratar de identificar, a nivel de microdato, la propensión real de ocultar que tiene cada persona y hogar, de manera que la corrección que se realice tenga principalmente esta información en consideración.

En la publicación "Una corrección de la distribución de ingresos de los hogares españoles procedente de la E.C.V. mediante técnicas de fusión estadística" (Callealta y Romero, 2013b) se presenta un primer intento de estos autores de allegar, al proceso de corrección, información de los hogares con respecto de su comportamiento económico, que pudiera dar algo de luz en este asunto. Y, como las ECV aportan cuantificaciones económicas de muy diversas facetas de los ingresos, pero casi exclusivamente de éstos (con la excepción de pagos para hipotecas), estos autores intentan allegar dicha información complementaria realizando una fusión estadística de las ECV con las EPF, con la esperanza de incorporar a los hogares de la primera información del gasto de esta última.

En Callealta y Romero (2013b), mediante procedimientos de fusión estadística basados en la disimilaridad de Gower, se complementa la información sobre rentas que aportan las ECV con otras dos variables sobre el consumo de los hogares: el gasto total en consumo del hogar (GASTOH) y el Índice de Cobertura de los gastos sobre los ingresos declarados por el hogar (IC).

El IC así fusionado con la información de las ECV permitió a los autores obtener una primera aproximación a la renta no declarada de cada hogar. De esta forma, si los gastos totales de un cierto hogar "h" (imputados por el proceso de fusión) superaban a los ingresos monetarios declarados por el mismo (cuando su IC era mayor que uno), procedían a estimarle la cantidad "Oh", presumiblemente ocultada o no declarada, como la proporción (IC-1) de las correspondientes rentas disponibles declaradas por el hogar en su conjunto.

Y como la posible subdeclaración de ingresos que se pudiera observar dentro de un hogar no tenía por qué ocurrir de forma uniforme entre sus miembros, y bajo la hipótesis de que más bien podrían depender de la composición y naturaleza de los ingresos que percibe cada miembro del hogar⁶³, los autores procedían a repartir dicha cantidad "Oh" entre sus miembros, asignando a cada miembro "m" de cada hogar "h" que percibe un tipo de renta "t", la cantidad " $O_{h,m}^t$ " calculada en función de la importancia relativa que tiene la infradeclaración de tal tipo de renta, " O^t ", en el contexto de las rentas infradeclaradas de forma agregada por los hogares en los tipos de rentas que su hogar percibe, así como de un cierto indicador " $y_{h,m}^t$ ", construido para cada miembro del hogar y tipo de renta "t", que estima una cierta propensión para éste a infradeclarar tal tipo de renta. Así, la infradeclaración " $O_{h,m}^t$ " se calculaba como:

$$O_{h,m}^t = O_h \frac{O^t I^t(h)}{\sum_t O^t I^t(h)} \frac{y_{h,m}^t}{\sum_m y_{h,m}^t}$$

siendo:

$$I^t(h) = \begin{cases} 1 & \text{si algún miembro del hogar h recibe el tipo de renta t} \\ 0 & \text{en otro caso} \end{cases}$$

Dicho indicador " $y_{h,m}^t$ " se puede obtener a partir de una o varias covariables, Y_1, Y_2, \dots, Y_k , presumiblemente correlacionadas con dicha propensión a la subdeclaración, todas ellas en la misma dirección, mediante el cálculo de la norma euclídea de su vector de coordenadas en el cubo multidimensional estandarizado de lados [0;1], a partir de la expresión:

⁶³ En Callealta y Romero, (2013a) se presentan los agregados de rentas de los hogares españoles derivados de las ECV para diferentes partidas según los tipos de renta percibidas, comparados con los correspondientes de la Cuenta de Rentas del Sector de Hogares de la CNE. Así, se observa que la subdeclaración apreciada en los hogares se comporta de forma muy diferente en las diversas partidas asociadas a los distintos tipos de renta.

$$y_{h,m}^{\dagger} = \sqrt{\sum_{p=1}^k \left(\frac{Y_{p,h,m} - \min_m(Y_{p,h,m})}{\max_m(Y_{p,h,m}) - \min_m(Y_{p,h,m})} \right)^2}$$

En relación a las partidas de rentas “†” consideradas candidatas a ser objeto de cierta reasignación individual de la rentas infradeclaradas globalmente por cada hogar, Callealta y Romero las limitan a las diversas partidas de ingresos percibidos, ya que son éstas las que se relacionan directamente con la mayor capacidad de gasto realizado por los miembros del hogar, dejando al margen e inalteradas las partidas de transferencias de rentas pagadas a otros hogares así como las de impuestos y cotizaciones sociales.

En consecuencia, consideran los cinco tipos de rentas siguientes, para las que indicamos entre paréntesis las covariables empleadas en la construcción del indicador de propensión a la infradeclaración en dicho tipo de renta: *rentas mixtas* (covariables: renta monetaria por cuenta propia declarada por cada miembro del hogar y renta propia per cápita declarada por su hogar de residencia), *rentas salariales* (covariables: renta personal monetaria por cuenta ajena declarada por cada miembro del hogar y el cociente entre ésta y las cotizaciones sociales per cápita declaradas en su hogar de pertenencia), *rentas de la propiedad* (covariables: rentas personales monetarias, tanto por cuenta propia como ajena, como indicadores del tamaño de su aportación al ahorro del hogar), *rentas procedentes de prestaciones sociales* (covariables: la suma de prestaciones sociales recibidas por la persona), y *rentas percibidas de otras transferencias corrientes* (covariables: las transferencias corrientes recibidas de otros hogares declaradas en el hogar, y el valor opuesto de la renta per cápita en el hogar, como indicador del suficiencia económica).

El proceso de reasignación de las rentas infradeclaradas estimadas para cada hogar entre sus miembros se ha realizado segmentadamente en cada una de las clases que el estudio considera (comunidades autónomas), ya que los autores observan diferencias entre ellas en las ocultaciones agregadas por tipo de renta⁶⁴.

2. Modelo propuesto para la estimación combinada de la DPR

2.1. Formulación inicial del modelo para la corrección de los microdatos

El modelo de partida que se propone para la estimación de una Distribución del tamaño de los ingresos en las ECV, congruente con los datos de CNE, y que se obtenga corrigiendo los casos de acuerdo con la información disponible para ellos sobre el grado de incongruencia de sus declaraciones y el grado de propensión a ocultar que de ella pueda derivarse, que no es sino el objetivo de este capítulo, se basa en el propuesto por esta

⁶⁴ Véanse Callealta y García (2011) y Callealta y Romero (2013a).

doctorando y su director de Tesis en la publicación "Sobre la congruencia de los datos de renta de las Encuestas de Condiciones de Vida y los de Contabilidad Nacional" (Callealta y Romero, 2013a).

Supongamos que en una clase "C" de personas, representada por una muestra de "n" casos "i" con factores de elevación "w_i", se detecta una ocultación agregada "O" para una cierta variable "Z" en la que los casos presentan valores declarados "z_i"⁶⁵.

Para la obtener la formulación matemática de la propuesta de corrección realizada en este trabajo, partiremos admitiendo las siguientes hipótesis:

- H₁: La ocultación "O_i" en una variable "Z", imputable a cada caso "i" de la clase "C", es del mismo signo que la ocultación total agregada en dicha clase, "O".

- H₂: La ocultación "O_i" en una variable "Z", imputable a cada caso "i" de la clase "C", es directamente proporcional al nivel del valor "x_i" presentado por "i" en alguna variable "X" (explicativa de la ocultación) con respecto del mínimo valor observado en la clase; de esta forma, se puede calcular $O_i = t_i \cdot (x_i - \min(X))$, siendo t_i el factor de proporcionalidad, que llamaremos "tasa de ocultación" con respecto a dicho nivel del valor observado, $t_i = O_i / (x_i - \min(X))$.

- H₃: La anterior tasa de ocultación t_i de un caso "i" depende linealmente del nivel presentado para el caso "i" en alguna variable "Y" que nos informa sobre su intensidad, con respecto de su menor valor observado; de forma que $t_i = a + b \cdot (y_i - \min(Y))$.

Admitidas estas hipótesis, el microdato corregido para el caso "i" en la variable Z tomará el nuevo valor "z_i+O_i". Y el problema se reduce, pues, a encontrar la ocultación O_i imputable a cada caso "i", sabiendo que su factor de elevación es "w_i", que presenta valores observados "x_i" e "y_i", respectivamente, para las variables X e Y, y que debe verificarse que:

$$O = \sum_{i=1}^n w_i O_i = \sum_{i=1}^n w_i t_i (x_i - \min X) = \sum_{i=1}^n w_i (a + b(y_i - \min Y))(x_i - \min X)$$

Notando $\tilde{x}_i = x_i - \min X$ e $\tilde{y}_i = y_i - \min Y$, la hipótesis H₁ implica, necesariamente, que "a" y b" deben tener el mismo signo que "O", pudiendo ser nulos.

Y de las hipótesis H₂ y H₃, se deduce que,

⁶⁵ Cuando nos refiramos a "O" como a una cierta "ocultación" en una cierta variable Z (partida de la Cuenta de Rentas de los hogares) agregada para una determinada clase de hogares, deberemos entender de forma general que con ella evaluamos la diferencia que, para la clase, se obtiene entre el valor que presenta la correspondiente partida en CNE y la que se calcularía a partir de los datos de las ECV. De esta forma, existe la posibilidad de encontrarnos en ella con ambos signos: cuando O>0, efectivamente habrá una verdadera ocultación de rentas en la clase; pero cuando O<0, estaremos refiriéndonos a una cierta sobredeclaración de este tipo de renta en esta clase.

$$O = \sum_{i=1}^n w_i \tilde{x}_i (a + b \tilde{y}_i) = a \sum_{i=1}^n w_i \tilde{x}_i + b \sum_{i=1}^n w_i \tilde{x}_i \tilde{y}_i \Leftrightarrow b = \frac{O - a \sum_{i=1}^n w_i \tilde{x}_i}{\sum_{i=1}^n w_i \tilde{x}_i \tilde{y}_i}$$

de donde, de forma general, la ocultación imputable al caso "i" será:

$$O_i = a \tilde{x}_i + \frac{O - a \sum_{i=1}^n w_i \tilde{x}_i}{\sum_{i=1}^n w_i \tilde{x}_i \tilde{y}_i} \tilde{x}_i \tilde{y}_i = \frac{\tilde{x}_i}{\sum_{i=1}^n w_i \tilde{x}_i \tilde{y}_i} \left\{ a \sum_{i=1}^n w_i \tilde{x}_i \tilde{y}_i + O \tilde{y}_i - a \tilde{y}_i \sum_{i=1}^n w_i \tilde{x}_i \right\}$$

Teniendo en cuenta lo anterior, el valor absoluto de la pendiente "b" será máximo (máxima progresividad en la ocultación) cuando a=0, en cuyo caso valdrá:

$$b_{\max} = \frac{O}{\sum_{i=1}^n w_i \tilde{x}_i \tilde{y}_i}$$

Y expresando "b" como una cierta proporción "k" de la máxima posible, "b_{max}", podremos expresar "a", en función de aquélla, como:

$$b = \frac{O - a \sum_{i=1}^n w_i \tilde{x}_i}{\sum_{i=1}^n w_i \tilde{x}_i \tilde{y}_i} = kb^* = k \frac{O}{\sum_{i=1}^n w_i \tilde{x}_i \tilde{y}_i} \Leftrightarrow a = \frac{(1-k)O}{\sum_{i=1}^n w_i \tilde{x}_i}$$

Consecuentemente, la ocultación imputada a la observación "i" en la variable Z y su correspondiente valor corregido en dicha variable serán, respectivamente:

$$O_i = \tilde{x}_i O \left\{ \frac{1-k}{\sum_{i=1}^n w_i \tilde{x}_i} + \frac{k \tilde{y}_i}{\sum_{i=1}^n w_i \tilde{x}_i \tilde{y}_i} \right\}, \quad Z_{\text{corr},i} = Z_i + O_i = Z_i + \tilde{x}_i O \left\{ \frac{1-k}{\sum_{i=1}^n w_i \tilde{x}_i} + \frac{k \tilde{y}_i}{\sum_{i=1}^n w_i \tilde{x}_i \tilde{y}_i} \right\}$$

2.2. Adaptación del modelo para que no dependa de la renta declarada

En Callealta y Romero (2013b), la ocultación imputable a cada caso en este modelo que acabamos de introducir se hace depender de forma proporcional de los niveles de las variables "X" directamente recogidas en las ECV y que, por tanto, son rentas declaradas de los distintos tipos, lo que produce correcciones mínimas de ocultación para las rentas declaradas más bajas. A su vez, esto contradice la teoría de que en las rentas bajas también se oculta (se elude la declaración) y hace plausible la idea de que las tasas de ocultación en estas zonas puedan ser relativamente elevadas.

Los autores allí tratan de mitigar este efecto haciendo que las tasas de ocultación t_i dependan de otras variables "Y" que, intuitivamente, deberían modular el efecto de la corrección con otra información complementaria; pero siguen siendo otros tipos de rentas declaradas las empleadas a tal fin, por lo que el problema de fondo persiste. Como ya hemos avanzado anteriormente, asumir una forma lineal para la tasa de ocultación dependiente de otras rentas análogamente declaradas no parece ser la opción más adecuada para llevar a cabo dicha corrección. Y, por añadidura, el método base del que partimos deja indeterminada la pendiente, o el coeficiente "k", de la función lineal con la que se modula la tasa de ocultación, abriéndose una amplia vía para la subjetividad del investigador.

En los siguientes apartados pretendemos ir configurando nuestra propuesta para tratar de dar una solución objetiva a estas debilidades del modelo base que acabamos de plantear.

Básicamente, nuestra propuesta consiste en modificar convenientemente dicho modelo base para no hacer depender la ocultación estimada de los hogares de sus rentas declaradas, sino de los indicios objetivos de la incongruencia que puedan observarse en sus declaraciones y que pudieran informarnos sobre su posible propensión a la ocultación. Las leyes económicas y/o empíricas que nos van a permitir encontrar tales posibles fuentes de incongruencia han sido ya expuestas en el capítulo anterior; y la medición del grado de las potenciales incongruencias que pudieran presentarse para la declaración de un determinado tipo de renta en cada hogar, se abordará en un próximo apartado de este mismo capítulo.

La primera consecuencia de modificar el método base en tal sentido es que hace más plausible la estimación de la cantidad ocultada por cada hogar a través de la cierta relación directa o positiva que pudiera haber entre ella y los valores que presenten los hogares en dichos indicadores de incongruencia. Y, por ello, la forma lineal de la propensión a ocultar cuando se presenta un determinado nivel de incongruencia en la declaración, que antes no era la más adecuada cuando dependía de las rentas declaradas, ahora se torna admisible cuando se la hace depender de tales indicadores de propensión a la ocultación.

Formulación del modelo

Supongamos que en una clase "C" de hogares, representada por una muestra de "n" casos "i" con factores de elevación " w_i ", se detecta una ocultación agregada "O" para una cierta variable "Z" en la que los casos presentan valores declarados " z_i ".

Para la obtener la formulación matemática de la corrección que proponemos, partiremos admitiendo las siguientes hipótesis:

- H₁: La ocultación " O_i " en una variable "Z", imputable a cada caso "i" de la clase "C", es del mismo signo que la ocultación total agregada en dicha clase, "O"

- H₂: La ocultación " O_i " en una variable "Z", imputable a cada caso "i" de la clase "C", es directamente proporcional al nivel de incongruencia " x_i " observado para el caso i en la variable Z mediante un indicador "X" que informa del nivel de incongruencia del valor " z_i " declarado para la variable

Z en relación a la Ocultación, considerado con respecto de su menor valor observado. De esta forma, se puede calcular $O_i = t_i \cdot (x_i - \min(X))$, siendo t_i su factor de proporcionalidad, $t_i = O_i / (x_i - \min(X))$.

- H₃: El factor de proporcionalidad t_i empleado para el cálculo de la ocultación "O_i" de un caso "i" se supondrá que varía linealmente con el nivel de incongruencia y_i presentado por el caso, con respecto de su menor valor observado, para algún indicador "Y" que nos informa sobre la propensión global a la ocultación del mismo. Así, el factor de proporcionalidad adopta la forma $t_i = a + b \cdot (y_i - \min(Y))$, intensificando la corrección realizada en base al indicador X, en función de la propensión a ocultar, en el agregado de todas sus rentas, que el caso muestre a partir del indicador Y.

A continuación, pasamos a exponer la metodología que proponemos para medir tales indicadores, a partir de la información adicional allegada por las fusiones estadísticas realizadas en el capítulo III de esta tesis entre las fuentes complementarias a las ECV aquí consideradas.

2.3. Metodología propuesta para la detección de observaciones incongruentes y la medida de su grado de incongruencia o propensión a la ocultación

En primer lugar, vamos a hacer un repaso de lo que entendemos cuando nos referimos a "casos concordantes" y "casos discordantes", conceptos que se emplean en las definiciones de medidas de asociación tan conocidas como pueden ser las Tau de Kendall o la Delta de Sommers, entre otras, utilizadas para la medición del grado de asociación entre variables cualitativas.

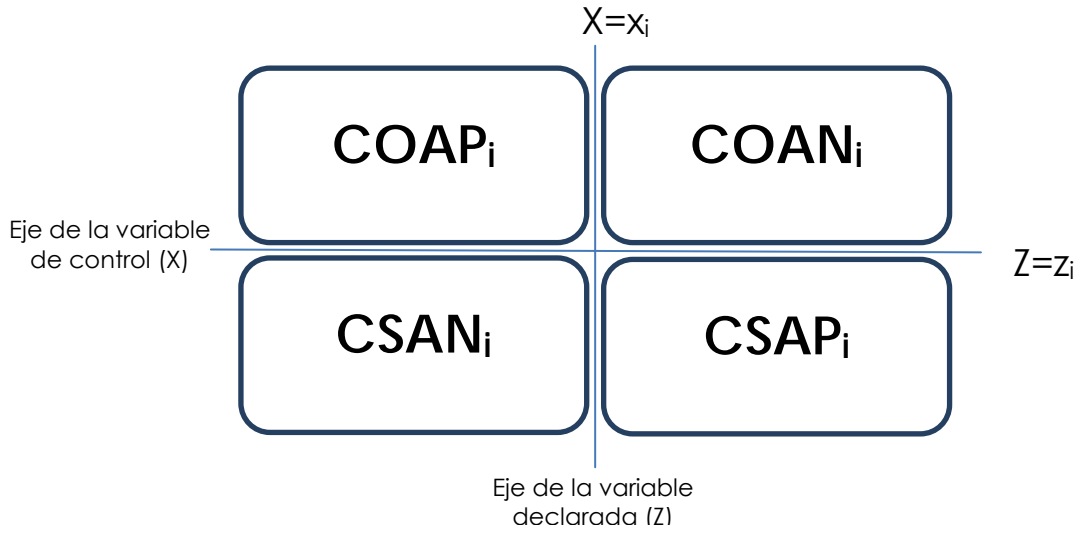
Supongamos que la variable Z es la variable de interés, a corregir por ocultación o sobredeclaración, y que X es una variable auxiliar que presenta una cierta asociación, positiva o negativa, con ella.

Decimos que dos pares de observaciones (x_i, z_i) y (x_j, z_j) son concordantes si $x_i < x_j$ y $z_i < z_j$, o bien $x_i > x_j$ y $z_i > z_j$. Como puede observarse, cada par de observaciones que sea concordante, nos informará de un cierto apoyo a algún tipo de asociación ordinal positiva entre las variables X y Z.

Análogamente, decimos que dos pares de observaciones (x_i, z_i) y (x_j, z_j) son discordantes si $x_i < x_j$ y $z_i > z_j$, o bien $x_i > x_j$ y $z_i < z_j$. De este modo, cada par de observaciones que sea discordante, informa de un cierto apoyo a algún tipo de asociación ordinal negativa entre las variables X y Z.

En la misma situación, y como punto de partida para definir nuestros indicadores de incongruencia o propensión a la ocultación, observemos que cada observación (x_i, z_i) divide el conjunto de observaciones en cuatro clases o grupos de casos, según se representa en la figura 4.1 que se muestra a continuación.

Figura 4.1
CLASES INDUCIDAS POR UNA RELACIÓN ORDINAL
PARA CADA CASO (x_i, z_i)



Fuente: Elaboración propia.

Cada una de estas clases estaría compuesta por las siguientes observaciones:

- Clase $COAP_i$ (Casos en situación de Ocultación y Asociación Positiva):

$$COAP_i = \left\{ (x_j, z_j) \mid x_i > x_j \wedge z_i \leq z_j ; j = 1, 2, \dots, n \right\} \quad \forall i = 1, 2, \dots, n$$

Las observaciones que conforman esta clase se muestran discordantes con la observación i -ésima de referencia de la clase informando de un cierto grado de ocultación en z_i con respecto de ellas, cuando entre Z y X existe una cierta asociación positiva. Obsérvese que, para todas las observaciones de esta clase, la observación i -ésima presenta un valor en la variable X mayor que todas ellas. Pero, en contra de lo esperado para la asociación positiva existente entre X y Z (a mayor valor de X , mayor valor de Z), la observación i -ésima de referencia presenta un valor en Z menor que todas ellas. Por ello, esta clase nos informará de una cierta propensión a la ocultación en la variable Z para la observación i -ésima, comparativamente al comportamiento observado en el resto de observaciones.

- Clase $COAN_i$ (Casos en situación de Ocultación y Asociación Negativa):

$$COAN_i = \left\{ (x_j, z_j) \mid x_i < x_j \wedge z_i \leq z_j ; j = 1, 2, \dots, n \right\} \quad \forall i = 1, 2, \dots, n$$

Las observaciones que conforman esta clase se muestran concordantes con la observación i -ésima de referencia de la clase, informando de un cierto grado de ocultación en z_i con respecto de ellas, cuando entre Z y X existe una cierta asociación negativa. Obsérvese que, para todas las observaciones de esta clase, la observación i -ésima presenta un valor en la variable X menor que todas ellas. Sin embargo, en contra de lo

esperado para la asociación negativa existente entre X y Z (a menor valor de X, mayor valor de Z), la observación i-ésima de referencia presenta también un valor en Z menor que todas ellas. Por ello, esta clase nos informará de una cierta propensión a la ocultación en la variable Z para la observación i-ésima, comparativamente al comportamiento observado en el resto de observaciones.

- Clase $CSAP_i$ (Casos en situación de Sobredeclaración y Asociación Positiva):

$$CSAP_i = \left\{ (x_j, z_j) \mid x_i < x_j \wedge z_i \geq z_j ; j = 1, 2, \dots, n \right\} \quad \forall i = 1, 2, \dots, n$$

Las observaciones que conforman esta clase se muestran discordantes con la observación i-ésima de referencia de la clase informando de un cierto grado de sobredeclaración en z_i con respecto de ellas, cuando entre Z y X existe una cierta asociación positiva. Obsérvese que, para todas las observaciones de esta clase, la observación i-ésima presenta un valor en la variable X menor que todas ellas. Pero, en contra de lo esperado para la asociación positiva existente entre X y Z (a menor valor de X, menor valor de Z), la observación i-ésima de referencia presenta un valor en Z mayor que todas ellas. Por esta razón, esta clase nos informará de una cierta propensión a la sobredeclaración en la variable Z para la observación i-ésima, comparativamente al comportamiento observado en el resto de observaciones.

- Clase $CSAN_i$ (Casos en situación de Sobredeclaración y Asociación Negativa):

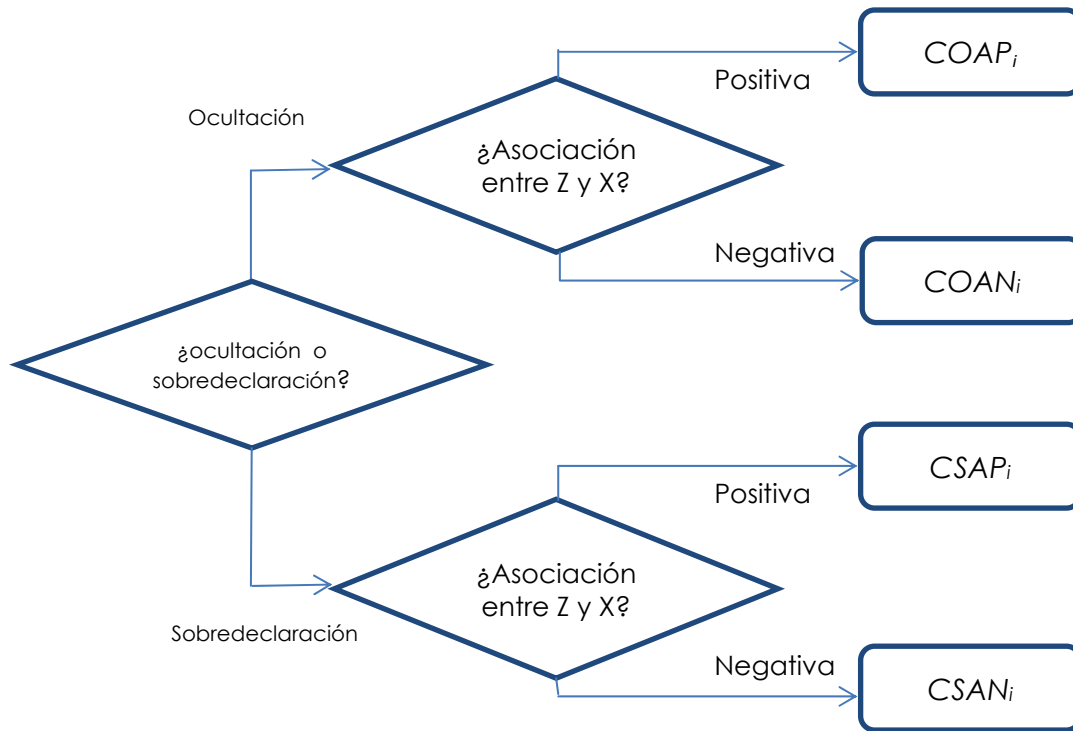
$$CSAN_i = \left\{ (x_j, z_j) \mid x_i > x_j \wedge z_i \geq z_j ; j = 1, 2, \dots, n \right\} \quad \forall i = 1, 2, \dots, n$$

Las observaciones que conforman esta clase se muestran concordantes con la observación i-ésima de referencia de la clase, informando de un cierto grado de sobredeclaración en z_i con respecto de ellas, cuando entre Z y X existe una cierta asociación negativa. Obsérvese que, para todas las observaciones de esta clase, la observación i-ésima presenta un valor en la variable X mayor que todas ellas. Sin embargo, en contra de lo esperado para la asociación negativa existente entre X y Z (a mayor valor de X, menor valor de Z), la observación i-ésima de referencia presenta también un valor en Z mayor que todas ellas. Por ello, esta clase nos informará de una cierta propensión a la sobredeclaración en la variable Z para la observación i-ésima, comparativamente al comportamiento observado en el resto de observaciones.

Así pues, dependiendo de si en una variable Z se observa generalizadamente el fenómeno de una cierta ocultación, o contrariamente el de una cierta subdeclaración, la consideración de alguna variable X asociada positiva o negativamente con aquélla nos permitirán identificar, para cada observación "i", el conjunto de casos que nos alertaría de una posible ocultación o subdeclaración en dicha observación "i".

En la siguiente figura (4.2) se refleja cuáles serían estas regiones en cada caso de interés:

Figura 4.2
**IDENTIFICACIÓN DE CLASES DE INTERÉS PARA LA
CONSTRUCCIÓN DE INDICADORES DE INCONGRUENCIA**



Fuente: Elaboración propia.

Consecuentemente, a partir de estas regiones, podremos derivar algunos indicadores que nos informen del grado de incongruencia del valor que presenta para la variable Z cada observación i-ésima con respecto de las demás de acuerdo con la información auxiliar que suministra otra variable X asociada con ella, a partir de las características geométricas de las correspondientes clases de interés señaladas, dependientes de la situación.

Así, podemos definir los siguientes:

- Indicador de tipo "Ic":

$$I_{C_i} = \frac{\sum_{j=1, j \neq i}^n Ind_{C_i}(x_j, z_j)}{n-1}, \text{ siendo } Ind_{C_i}(x_j, z_j) = \begin{cases} 1 & \text{si } (x_j, z_j) \in C_i \\ 0 & \text{si } (x_j, z_j) \notin C_i \end{cases}$$

Este indicador nos dará una idea del grado de apoyo que el conjunto de observaciones da a la creencia de la incongruencia de z_i , medido como la proporción del total de casos comparables con la observación

i-ésima de referencia que se encuentran en la clase "c" de interés, según el tipo de relación existente entre X y Z.

- Indicador de tipo "Dc":

$$D_{c_i} = \frac{\sum_{j \in C_i}^n (x_j - x_i)^2 + (z_j - z_i)^2}{n-1}$$

Este indicador nos informará acerca de la magnitud de la incongruencia detectada para el valor z_i , medida a partir del cálculo de la lejanía media de la observación i-ésima de referencia a los casos de la clase "c" de interés, en términos del cuadrado de la distancia euclídea, en base al tipo de relación que exista entre X y Z.

- Indicador de tipo "Jc":

$$J_{c_i} = \frac{\sum_{j \in C_i}^n |x_j - x_i| \cdot |z_j - z_i|}{n-1}$$

Este indicador nos dará una idea de la magnitud de la incongruencia detectada para el valor z_i , medida a partir de la magnitud de la separación entre la observación i-ésima de referencia y la clase "c" de interés, según el tipo de relación entre X y Z, calculada como el área media de los rectángulos que determinan el caso i-ésimo con cada uno de las observaciones de la clase "c" consideradas en el vértice opuesto.

En el Anexo 3 se adjunta el programa informático mediante el que podemos calcular estas medidas de incongruencia (allí referidas como "discordancias") y que aplicaremos a las partidas de nuestra Cuenta de Rentas de los Hogares en el apartado 4 de este capítulo.

Dado que estos indicadores miden diferentes matices sobre la potencial incongruencia que presenta una observación para una variable Z con respecto del resto de observaciones y, de acuerdo con la información suministrada por una variable auxiliar X asociada, podremos extraer otros indicadores adicionales, por el mismo método, si disponemos de otras variables diferentes también asociadas con Z.

Como consecuencia de este proceso, cada observación dispondrá de un número variable de indicadores de incongruencia con respecto de las demás observaciones. Y como todos estos indicadores tienen en común el hecho de informar del grado de incongruencia del valor declarado, en relación a la ocultación o sobredeclaración agregada, según el caso, en la correspondiente variable, podremos sintetizar la esencia de estos indicadores a través de su primera componente principal. Por ello, y de acuerdo con esto, proponemos caracterizar finalmente el grado de incongruencia de cada observación z_i como el valor obtenido para ella a partir de la primera componente principal de todos los indicadores calculados a tal fin.

2.4. Estimación de mínimos cuadrados de la propensión a la ocultación indeterminada, k

Una vez calculados los pertinentes indicadores de incongruencia mediante la metodología descrita previamente, y admitidas las hipótesis H_1 a H_3 del modelo propuesto para la corrección de las ocultaciones, el microdato corregido para el caso "i" en la variable Z tomará el nuevo valor " $z_i + O_i$ ", siendo la cantidad de ocultación estimada para dicho caso "i", análogamente a como se ha presentado para el modelo base, y notando como allí $\tilde{x}_i = x_i - \min_{j=1,2,\dots,n} \{x_j\}$ e $\tilde{y}_i = y_i - \min_{j=1,2,\dots,n} \{y_j\}$:

$$O_i = t_i (x_i - \min X) = (a + b(y_i - \min Y))(x_i - \min X)$$

$$\text{cumpliéndose que: } O = \sum_{i=1}^n w_i O_i$$

La hipótesis H_1 implicaría, necesariamente, que los signos de "a" y "b" deben ser los mismos que el de "O", pudiendo ser nulos cuando la ocultación agregada O sea nula⁶⁶.

Análogamente a como hemos visto ya en el modelo base, de las hipótesis H_2 y H_3 , se deducen las expresiones de las correcciones que el modelo haría sobre cada declaración, siendo:

$$O_i = \tilde{x}_i O \left\{ \frac{1-k}{\sum_{i=1}^n w_i \tilde{x}_i} + \frac{k \tilde{y}_i}{\sum_{i=1}^n w_i \tilde{x}_i \tilde{y}_i} \right\}, \quad z_{\text{corr},i} = z_i + O_i = z_i + \tilde{x}_i O \left\{ \frac{1-k}{\sum_{i=1}^n w_i \tilde{x}_i} + \frac{k \tilde{y}_i}{\sum_{i=1}^n w_i \tilde{x}_i \tilde{y}_i} \right\}$$

Y de aquí, dada una cierta ocultación agregada O para las declaraciones de la variable de interés, las situaciones extremas para estas correcciones se darían cuando el factor de proporcionalidad para la corrección fuese constante y no presentara ninguna progresividad con el nivel de incongruencia del caso para la variable a corregir (lo que ocurriría cuando $k=0$) o cuando el factor de proporcionalidad presentara la máxima progresividad con el nivel de incongruencia del caso para la variable a corregir (lo que ocurriría cuando $k=1$).

Se podría eliminar la indeterminación de la pendiente "k" en esta solución, si se dispusiera de información que ligara el cómo las personas y hogares tendieran a comportarse con respecto al fenómeno de la ocultación.

Bien sabido es que esta información, que podría derivarse de los resultados de la labor de inspección fiscal que se realiza por parte del estado rutinariamente, ya sea para proteger el carácter privado de los datos o bien

⁶⁶ Si $O > 0$, la ocultación O_i estimada para el caso "i" será una cantidad positiva, resultante de multiplicar el grado de incongruencia de la observación por una tasa positiva de propensión a la ocultación y que crece linealmente con el grado de incongruencia de la observación. Por el contrario, si $O < 0$, la sobredeclaración estimada O_i será una cantidad negativa, resultante de multiplicar el grado de incongruencia de la observación por una tasa negativa de propensión a la sobredeclaración y que decrece linealmente con el grado de incongruencia de la observación.

para no menoscabar la eficacia de las futuras labores de inspección, no trasciende a los investigadores. Sin embargo, como veremos en un próximo apartado, la comparación de los microdatos obtenidos para las ECV en las dos bases publicadas (los datos de la base 2004, como expresión de lo inicialmente declarado por las familias; y los de la base 2013, como expresión resultante de una cierta corrección de las ocultaciones de aquellas declaraciones) posibilitan la realización del análisis de las cantidades ocultadas por las familias y el grado de dependencia de éstas con relación a sus indicadores de incongruencia, una vez calculados éstos.

Así, supongamos que la muestra se puede segmentar en H grupos de hogares. Podríamos deducir de las expresiones anteriores que:

$$O_i = \tilde{x}_i O \left\{ \frac{1-k}{\sum_{i=1}^n w_i \tilde{x}_i} + \frac{k \tilde{y}_i}{\sum_{i=1}^n w_i \tilde{x}_i \tilde{y}_i} \right\} \Rightarrow O_h^* = \sum_{i=1}^{n_h} \tilde{x}_i O \left\{ \frac{1-k}{\sum_{h=1}^H \sum_{i=1}^{n_h} w_i \tilde{x}_i} + \frac{k \tilde{y}_i}{\sum_{h=1}^H \sum_{i=1}^{n_h} w_i \tilde{x}_i \tilde{y}_i} \right\} \Leftrightarrow$$

$$\Leftrightarrow O_h^* = (1-k) O \frac{\sum_{i=1}^{n_h} \tilde{x}_i}{\sum_{h=1}^H \sum_{i=1}^{n_h} w_i \tilde{x}_i} + k O \frac{\sum_{i=1}^{n_h} \tilde{x}_i \tilde{y}_i}{\sum_{h=1}^H \sum_{i=1}^{n_h} w_i \tilde{x}_i \tilde{y}_i}$$

Y supuestas conocidas las cantidades O_h^* ocultadas por cada uno de estos grupos, "h", compuesto por n_h hogares de la muestra inicial (cada uno con un peso w_i en la muestra total), entonces el error de estimación de la cantidad ocultada por el grupo "h" cometida por nuestro modelo sería:

$$e_h = O_h^* - O_h = \left(O_h^* - O \frac{\sum_{i=1}^{n_h} \tilde{x}_i}{\sum_{h=1}^H \sum_{i=1}^{n_h} w_i \tilde{x}_i} \right) - k O \left(\frac{\sum_{i=1}^{n_h} \tilde{x}_i \tilde{y}_i}{\sum_{h=1}^H \sum_{i=1}^{n_h} w_i \tilde{x}_i \tilde{y}_i} - \frac{\sum_{i=1}^{n_h} \tilde{x}_i}{\sum_{h=1}^H \sum_{i=1}^{n_h} w_i \tilde{x}_i} \right)$$

De modo que, en el caso de que se conozcan dichas cantidades O_h^* podemos realizar la estimación mínimo cuadrática ocultada por cada grupo, del parámetro "k" inicialmente indeterminado en el modelo.

Procediendo de este modo, y llamando w_h^* a la suma de pesos individuales de los componentes del grupo h-ésimo:

$$w_h^* = \sum_{i=1}^{n_h} w_i$$

así como ψ a la suma de los cuadrados de los correspondientes residuos de los grupos de hogares:

$$\psi = \sum_{h=1}^H w_h^* e_h^2$$

deduciríamos que:

$$\begin{aligned}\psi &= \sum_{h=1}^H w_h^* e_h^2 \Rightarrow \frac{d\psi}{dk} = 2 \sum_{h=1}^H w_h^* e_h \frac{de_h}{dk} = 0 \Leftrightarrow \\ &\Leftrightarrow \sum_{h=1}^H w_h^* e_h \left(\frac{\sum_{i=1}^{n_h} \tilde{x}_i \tilde{y}_i}{\sum_{h=1}^H \sum_{i=1}^{n_h} w_i \tilde{x}_i \tilde{y}_i} - \frac{\sum_{i=1}^{n_h} \tilde{x}_i}{\sum_{h=1}^H \sum_{i=1}^{n_h} w_i \tilde{x}_i} \right) = 0\end{aligned}$$

de donde:

$$\begin{aligned}\sum_{h=1}^H w_h^* \left(O_h^* - O \frac{\sum_{i=1}^{n_h} \tilde{x}_i}{\sum_{h=1}^H \sum_{i=1}^{n_h} w_i \tilde{x}_i} \right) \left(\frac{\sum_{i=1}^{n_h} \tilde{x}_i \tilde{y}_i}{\sum_{h=1}^H \sum_{i=1}^{n_h} w_i \tilde{x}_i \tilde{y}_i} - \frac{\sum_{i=1}^{n_h} \tilde{x}_i}{\sum_{h=1}^H \sum_{i=1}^{n_h} w_i \tilde{x}_i} \right) &= \\ &= k O \sum_{h=1}^H w_h^* \left(\frac{\sum_{i=1}^{n_h} \tilde{x}_i \tilde{y}_i}{\sum_{h=1}^H \sum_{i=1}^{n_h} w_i \tilde{x}_i \tilde{y}_i} - \frac{\sum_{i=1}^{n_h} \tilde{x}_i}{\sum_{h=1}^H \sum_{i=1}^{n_h} w_i \tilde{x}_i} \right)^2\end{aligned}$$

pudiendo despejarse el valor estimado para "k" como:

$$k = \frac{\sum_{h=1}^H w_h^* \left(O_h^* - O \frac{\sum_{i=1}^{n_h} \tilde{x}_i}{\sum_{h=1}^H \sum_{i=1}^{n_h} w_i \tilde{x}_i} \right) \left(\frac{\sum_{i=1}^{n_h} \tilde{x}_i \tilde{y}_i}{\sum_{h=1}^H \sum_{i=1}^{n_h} w_i \tilde{x}_i \tilde{y}_i} - \frac{\sum_{i=1}^{n_h} \tilde{x}_i}{\sum_{h=1}^H \sum_{i=1}^{n_h} w_i \tilde{x}_i} \right)}{O \sum_{h=1}^H w_h^* \left(\frac{\sum_{i=1}^{n_h} \tilde{x}_i \tilde{y}_i}{\sum_{h=1}^H \sum_{i=1}^{n_h} w_i \tilde{x}_i \tilde{y}_i} - \frac{\sum_{i=1}^{n_h} \tilde{x}_i}{\sum_{h=1}^H \sum_{i=1}^{n_h} w_i \tilde{x}_i} \right)^2}$$

que es efectivamente la solución de mínimos cuadrados, ya que:

$$\frac{d^2\psi}{dk^2} = 2 \sum_{h=1}^H w_h^* \left(\left(\frac{de_h}{dk} \right)^2 + e_h \frac{d^2e_h}{dk^2} \right) = \sum_{h=1}^H w_h^* \left(\frac{de_h}{dk} \right)^2 \geq 0$$

En todo caso, la implementación práctica de este modelo requerirá, en consecuencia, información adicional para construir los indicadores de incongruencia propuestos en el apartado 2.3, así como información complementaria sobre ocultación de rentas que permita la estimación efectiva, para cada clase, del parámetro "k" inicialmente indeterminado y que ahora podríamos determinar de forma objetiva.

La información necesaria para construir los indicadores de incongruencia nos la proporcionarán las fuentes fusionadas en el capítulo III de este trabajo. Sin embargo, la información sobre la conducta que rige el fenómeno de la ocultación en las familias no se encuentra tan fácilmente como para disponer de ella en la investigación. Probablemente, los organismos de control correspondientes (Ministerio de Hacienda, organismos autonómicos con competencias fiscales, etc.) deban tener información muy rica al respecto,

procedente de sus actos de inspección; pero no es pública y se guarda celosamente.

La posibilidad que hemos encontrado para ello dentro de este trabajo, la presentamos en el apartado siguiente y proviene de la comparación de los resultados de las dos bases distintas de las ECV que el INE ha publicado para los años 2008 a 2012 (con años de referencia de los ingresos recogidos de 2007 a 2011, respectivamente).

3. Las ECV (base 2004) y su corrección con el uso de registros administrativos en las ECV (base 2013): una vía para el de análisis de la ocultación de los hogares

El fenómeno de la infradeclaración y el desajuste existente entre las encuestas y la Contabilidad Nacional no es un hecho privativo de nuestro país, sino que se manifiesta de forma similar en otros (Hurst y Pugsley, 2014; Susin, 2003; Wang y Woo, 2011; Dunbar y Fu, 2015; Kukk y Staehr, 2013; Paulus, 2015), dando pie a una reciente corriente internacional que trata de ver la posibilidad de paliar tal efecto mediante la complementación de las encuestas con el uso de registros administrativos (Cristia y Schwabish, 2009; Dauphin, 2009; Consolini, 2009). Así las cosas y conscientes de esta necesidad, la Oficina Estadística de la Unión Europea recopila experiencias de diversos países (Eurostat, 2013b) y establece las directrices de la “ESS Vision 2020”⁶⁷.

En esta línea y consecuentemente, también el INE comenzó en 2013 a utilizar registros administrativos, tratando de mejorar la calidad de los resultados de estas encuestas (Méndez y Vega, 2011 y 2014; INE, 2014).

El INE mantiene actualmente convenios de colaboración en materia estadística con la Agencia Estatal de Administración Tributaria, la Seguridad Social, la Hacienda Tributaria de Navarra y la Diputación Foral de Bizkaia. Este hecho está permitiendo un intercambio eficiente de información tal que, como hemos dicho anteriormente, ha hecho posible que las Encuestas de Condiciones de Vida, desde la ola de 2013, hayan incorporado un cambio metodológico que, básicamente, consiste en que las variables de ingresos se recogen prioritariamente a través de los ficheros administrativos proporcionados por aquellos organismos, a partir del NIF de las personas pertenecientes a la muestra.

Sin embargo, como los registros administrativos por sí solos se muestran insuficientes para determinar todas las rentas de los hogares y sus integrantes (ya que presentan problemas de cobertura geográfica o de cobertura de todos los perceptores de rentas, entre otros), la nueva metodología de las ECV, además de considerar los registros administrativos, mantiene en los cuestionarios las preguntas sobre ingresos que son necesarias para complementar éstos a la hora de construir las definitivas variables de ingresos

⁶⁷ Los miembros del Sistema Estadístico Europeo (ESS) definieron su estrategia de largo plazo y su programa de modernización como “Vision ESS 2020” y lo adoptaron en mayo de 2014. El refuerzo de la calidad fue identificado como uno de los cinco ámbitos clave que requieren una actuación común en el futuro (Eurostat, 2014b).

de las personas y hogares. Pero, en todo caso, hay muchas esperanzas puestas en que esta nueva metodología mejore la estimación de la verdadera distribución de la renta de los hogares.

En los siguientes apartados trataremos de explotar la presumible mejoría que aporta esta nueva metodología a las estimaciones que sobre ingresos proporcionan ahora las nuevas ECV (en adelante, ECV base 2013), centrándonos, especialmente, en el análisis de las correcciones que implícitamente está realizando para cada partida de renta, en relación con lo que estimaba la antigua metodología (en adelante, ECV base 2004).

Un primer análisis de cómo ha afectado la nueva metodología a la distribución personal de la renta extraída de las ECV ha sido abordado en Callealta y Romero, 2016. Allí, como primeras conclusiones generales de la comparación realizada entre las ECV base 2013 y las ECV base 2004 a nivel agregado, se afirmaba que la utilización de los registros administrativos consigue elevar los niveles de renta estimados para las personas y hogares, aproximando mejor las cifras agregadas de la CNE; y que ello sucede no sólo para la renta disponible agregada de los hogares sino también para las diversas partidas de renta que la integran, con la excepción de "Otras transferencias corrientes", que siguen siendo una asignatura pendiente para estas encuestas.

Además, la nueva metodología parece corregir sustancialmente los aparentes sesgos, o indicios contrarios a lo que se pudiera pensar, que se daban con la antigua metodología (por ejemplo, tasas de pobreza presumiblemente demasiado altas en los hogares en los que vivían personas que eran *empresarios agrarios y no agrarios, con y sin asalariados*, entre otros hechos un tanto sorprendentes a primera vista).

Y, aunque la nueva metodología sigue subdeclarando los niveles de renta agregada que la CNE obtiene, lo hace en menor medida y de forma más homogénea para las distintas partidas de lo que lo hacía la metodología anterior.

Así pues, la incorporación de los registros administrativos a las ECV, desde la perspectiva de los ingresos, parece haber sido muy positiva, aunque los autores sugieren que, para estimar una distribución más realista de las rentas reales de las personas y hogares es importante seguir trabajando en la línea de completar los vacíos informativos que aún tienen las nuevas encuestas (partidas que aún se encuentran mal recogidas); y, paralelamente, en la de investigar y describir las características y conducta de las personas que ocultan rentas que debían declarar.

En este sentido, el cambio de metodología de las ECV nos brinda una de las pocas oportunidades objetivas de indagar en el fenómeno de la subdeclaración. Por un lado, las ECV en la base 2004 nos proporcionan los datos inicialmente declarados por los encuestados (en algunos casos, imputados a cada uno de ellos mediante procedimientos para la corrección de la falta de respuesta, que es otra forma de ocultar). Por otro, las ECV en la base 2013 ofrecen la información corregida de las declaraciones de estos mismos encuestados, una vez que fueron confrontados con los registros administrativos.

En consecuencia, y a nivel de microdato, parece sencillamente posible calcular la cantidad ocultada por cada persona y hogar en cada tipo de ingreso, y confrontarla con sus características para tratar de profundizar en el estudio de los comportamientos de los que subdeclaran.

Sin embargo, y aunque finalmente ha sido posible realizar esta comparación básica de los correspondientes microdatos para el análisis de la ocultación de personas y hogares, el procesamiento de los datos no ha sido nada sencillo. El motivo de tal dificultad no ha sido otro que los códigos de identificación de personas y hogares en los archivos de microdatos de las ECV facilitados por el INE son diferentes en cada una de las bases; han sido completamente cambiados.

3.1. Identificación de los hogares de las ECV en las bases 2004 y 2013 mediante procedimientos de fusión estadística

La tarea de identificar las personas y hogares en las dos bases de las encuestas se ha llevado a cabo, nuevamente, mediante técnicas de fusión estadística; y describimos, a continuación, el método de fusión empleado con tal objetivo.

En primer lugar, se han construido sendas bases de datos personales, una para la ECV base 2004 y otra para la ECV base 2013, en las que cada registro de cada una de las personas se complementa con las variables de los hogares al que pertenece. Así, se dispone en cada registro de toda la información de la persona acompañada de la de su entorno familiar.

Con la base de datos de las ECV base 2004 como archivo maestro y la base de datos de las ECV base 2013 como archivo de consulta o control, se ha empleado una técnica de fusión estadística para identificar qué personas del archivo de control se corresponde con cuál del archivo maestro, buscando la igualdad de sus características personales y de las de su hogar de pertenencia; y, una vez identificados, terminar allegando al archivo maestro (base 2004) los correspondientes códigos de identificación de cada persona y de su hogar de pertenencia en el archivo de control (base 2013) como información fusionada.

Las características personales y de los hogares empleadas en este proceso son las relacionadas en los cuadros 4.1.a y 4.1.b.

Cuadro 4.1.a
**VARIABLES COMUNES DE CONTROL PARA LA FUSIÓN DE
 LAS BASES DE LAS ECV (PERSONAS)**

<u>Variables de las personas</u>
Numero de orden de la persona en el cuestionario del hogar
Ponderación transversal de la persona
Ponderación transversal del mayor de 16 años
Sexo de la Persona
Edad de la Persona
Estado civil de la persona
Nivel de Estudios de la Persona
Nacionalidad de la persona
Situación de la persona en relación con la actividad
Situación profesional de la persona
Ocupación de la persona en el último empleo
Categoría socio-económica de la Persona
Actividad del establecimiento del que depende o dependía laboralmente
Tipo de contrato de la persona
¿Es la persona dependiente?

Fuente: Elaboración propia.

Cuadro 4.1.b
**VARIABLES COMUNES DE CONTROL PARA LA FUSIÓN DE
 LAS BASES DE LAS ECV (HOGARES)**

<u>Variables de los hogares de las personas</u>
Ponderación transversal del hogar
Comunidad Autónoma de ubicación del hogar
Hábitat del Hogar
Tipo de vivienda
Régimen de tenencia de la vivienda
Número de habitaciones de la vivienda
Número de miembros del hogar
Número de miembros del hogar mayores de 16 años
Número de miembros del hogar mayores de 14 años
Tipo de hogar (clasificación sociológica)
Valor de Alquiler imputado Vivienda Principal

Fuente: Elaboración propia.

Análogamente a como hemos hecho en las fusiones llevadas a cabo en el capítulo III, para evaluar la afinidad de las personas en la ECV base 2004 y en la ECVbase 2013 hemos empleado como medida de proximidad la medida de disimilitud basada en la medida de disimilaridad de Gower, tal como allí se ha definido.

Como mecanismo de enlace para la fusión, se ha empleado el método del vecino más próximo, consistente en seleccionar como información a fusionar con la de cada persona de la ECV base 2004, la información de la persona más próxima (es decir, la persona que menor disimilitud calculada presente con aquella). Una vez determinada ésta, se captura su código de identificación en la base 2013; y, a partir de éste, se dispone de todas las variables que hayan sido modificadas en la base 2013.

Y para que el proceso de cálculo fuera más eficiente, hemos segmentado el proceso de fusión para cada región y cada ola. Así, para cada una de las

personas en la ECV base 2004 hemos calculado la disimilitud que presentaba con cada una de las personas de la misma región que se recogían en la ECV (base 2013) de la misma ola, utilizando para ello la anterior medida de proximidad sobre las variables comunes que se muestran en los cuadros 4.1.a y 4.1.b.

En el peor de los casos, en las ECV de 2008, 96 casos de personas han tenido que ser identificados manualmente (el 0,267% de los casos). En el resto, el proceso funcionó aún mejor requiriendo asignar manualmente a sólo 7, 2, 6 y 1 casos, respectivamente, para las ECV de 2009 a 2012, representando en cada año el 0,019%, 0,005%, 0,017% y el 0,003% de los casos.

3.2. Análisis de las ocultaciones desveladas por las ECV

Vamos a realizar algunas consideraciones previas sobre la tipología de los casos que podrían incurrir en ocultación de sus rentas, desde la pura óptica de la recogida de datos.

Dado un determinado tipo de ingresos (supongamos por ejemplo los derivados del alquiler de terrenos), el fenómeno de la ocultación se centraría, exclusivamente, en aquellos casos que podrían percibir este tipo de rentas. Llamaremos a este conjunto subpoblación de "afectados" por una potencial ocultación en este tipo de renta.

En la situación que nos ofrece el cambio de metodología de las ECV, tenemos información suficiente sobre las dos dimensiones que nos permiten determinar el conjunto de "afectados": la primera dimensión, que llamaremos dimensión "declaración", nos informa de qué unidades declararon dicho tipo de renta y en qué cuantía lo hicieron; la segunda dimensión, que llamaremos dimensión "ocultación", nos informa de qué unidades vieron corregidas su declaración (o su no declaración) y en qué cantidad.

Considerando estas dos dimensiones, podemos elaborar el siguiente cuadro (4.2), en el que se identifican los tipos de casos que podemos encontrar ante el fenómeno de la ocultación.

Cuadro 4.2
TIPOLOGÍA DE LOS CASOS AFECTADOS POR EL FENÓMENO DE LA OCULTACIÓN EN UN DETERMINADO TIPO DE RENTA

	Corrección <0 (sobredeclara)	No corrección	Corrección >0 (oculta)
Declaran <0	Afectado	Afectado	Afectado
No Declara	Afectado	No Afectado	Afectado
Declaran >0	Afectado	Afectado	Afectado

Fuente: Elaboración propia.

A partir de la información recogida en este cuadro, podemos determinar el conjunto de "afectados" para el estudio de la potencial ocultación en un determinado tipo de renta, como aquel colectivo formado por los casos que declararon alguna cantidad en dicho tipo de renta (positiva o negativa, hayan declarado correctamente o no), junto a los que no habiendo declarado cantidad alguna en ella, vieron corregida su no declaración (ocultaron o sobredeclararon la totalidad de sus ingresos).

El resto de casos, los que ni declararon en dicho tipo de renta ni se les realizó ninguna corrección, quedarían clasificados como "no afectados", presumiblemente porque no perciben este tipo de renta.

Por otra parte, como veremos en los cuadros que iremos viendo a lo largo de esta sección, hemos observado que la nueva metodología corrige registros no sólo al alza, sino que también corrige a la baja ciertas rentas declaradas. En el primer caso, se observan indicios inequívocos de ocultación manifiesta; en el segundo, en el que se han "sobredeclarado" los ingresos en cuestión, las ECV consideran lo declarado como erróneo y hacen valer la fuente de los registros administrativos como la válida, detrayendo de lo declarado la cantidad pertinente. Este efecto de "sobredeclaración" con relación a los registros administrativos podría deberse a varios motivos: posiblemente, a algún error en la declaración (quizás efecto "olvido" o "redondeo"); o quizás a una imputación excesivamente alta del INE en las ECV base 2004 ante la falta de respuesta.

Pensando en que el fenómeno de la ocultación de rentas se da también en los registros administrativos, ¿no podría interpretarse la falta de respuesta como un indicio de una cierta intención de ocultar en la encuesta? ¿no podría interpretarse la sobredeclaración como un indicio de una cierta intención de ocultar en los registros administrativos? Mientras haya diferencias entre lo declarado agregadamente en las encuestas con respecto de las Cuentas Nacionales, aún después de haber sido corregidas con los registros administrativos, creemos que esta información no debe despreciarse; por lo que desde el punto de vista de la dimensión "declaración" distinguiremos, cuando sea posible, las correcciones realizadas por sobredeclaración de las realizadas por ocultación.

Las diferencias entre los valores corregidos (registrados en los archivos de microdatos de la base 2013) y los valores declarados (registrados en los archivos de microdatos de la base 2004) recogen agregadamente tanto los efectos de la ocultación propiamente dicha, como de las infra y sobre imputaciones realizadas por los modelos que trataban la falta parcial o total de respuesta. Por tanto, debemos ser conscientes de que los siguientes cuadros que iremos presentando, mostrarán este efecto agregado, no habiendo sido posible en este trabajo aislar una potencial ocultación real de una infraimputación, o una potencial sobredeclaración real de una sobreimputación, producidas en el proceso de tratamiento de la no respuesta.

A continuación, presentamos los resultados obtenidos para cada una de las partidas de renta que hemos podido aproximar a partir de las ECV, sin emplear ninguna otra fuente externa para ello⁶⁸. Comenzamos analizando los

⁶⁸ Algunas partidas no han podido ser calculadas, ni siquiera aproximadamente, de forma directa a partir de los datos de las ECV, bien porque no se recogía su información, como es el caso de las D_44_r - Rentas de la

efectos sobre las partidas localizadas en la ECV que se aproximan a las de la Cuenta de Asignación de la Renta Primaria.

Efecto sobre el Excedente de explotación Neto

El cuadro 4.3 recoge el análisis del Excedente de Explotación Neto de las familias, cuya principal fuente es el Alquiler imputado a la vivienda principal de sus propietarios. Este hecho explica que el colectivo "afectado" para esta renta sea tan elevado (mayor del 80%).

Sin embargo, la corrección agregada realizada por la base 2013 ha sido muy pequeña en esta partida, no superando el 0,71% de la ocultación observada en la base 2004, en ninguno de los dos años analizados.

El porcentaje de hogares corregidos de entre el total de los que poseen este tipo de rentas entre la base 2004 y la de 2013 no supera en ninguno de los dos años el 2,1%. En cualquier caso, dichas correcciones se dieron tanto para elevar una cierta infradeclaración, como para rebajar una cierta sobredeclaración. Se observa poca propensión a ocultar este tipo de rentas cuando el valor declarado fue positivo (inferior al 0,6% en ambos años). Y cuando fue negativo, se observa una propensión mayor a ocultar en las declaraciones negativas, con unas ocultaciones medias en torno a los 1.750-2.000 euros.

propiedad atribuidas a los asegurados (recursos), D_61_r - Cotizaciones sociales (recursos), D_62_e - Prestaciones sociales distintas de transferencias sociales en especie (empleos), D_71_e - Primas netas de seguro no vida, D_72_r - Indemnizaciones de seguro no vida (recursos); bien porque se recogían agregadamente con otras partidas, como es el caso de la partida D_45_re - Rentas de la tierra (recursos-empleos), que se recogen agregadamente con los ingresos por el alquiler de propiedades que en CNE se incluyen en las Rentas Mixtas, o de las partidas D_51_e - Impuestos sobre la renta y D_61_e - Cotizaciones Sociales, que se recogen juntas en las ECV. Por su parte, la partida D_59_e - Otros impuestos corrientes, que se puede construir con la variable Impuestos sobre el patrimonio de la ECV no se ha analizado, ya que en uno de los años de estudio (2011) no se encuentra disponible en esta encuesta. El motivo fue que el Impuesto sobre el patrimonio se encontraba derogado al inicio de la recogida de los datos de la ECV y, aunque al final se volvió a restablecer en Septiembre de 2011, ya no dio tiempo recogerlo (Real Decreto-ley 13/2011, de 16 de septiembre, por el que se restablece el Impuesto sobre el Patrimonio, con carácter temporal).

Cuadro 4.3
EFFECTO SOBRE EL EXCEDENTE DE EXPLOTACIÓN NETO

B_2n	2008	2011
CNE (Residentes)	31.734.000.000	41.780.000.000
ECV (base 2004)	60.192.213.232	64.148.391.715
Ocultación	-28.458.213.232	-22.368.391.715
% Ocultación con respecto de CNE	-89,68%	-53,54%
ECV (base 2013) -€-	59.990.047.196	64.088.350.135
Ocultación aflorada	234.425.984	124.927.060
<i>Sobredeclaración aflorada</i>	<i>-436.592.019</i>	<i>-184.968.640</i>
Corrección neta	-202.166.036	-60.041.579
% Corrección en esta variable / Su ocultación (b2004)	0,71%	0,27%
Hogares de la muestra (con mayores de 16 años)	13.360	12.714
% Hogares afectados de la muestra	82,75%	83,41%
% de Hogares afectados corregidos	2,10%	1,40%
% Hogares afectados que ocultan	0,89%	0,58%
% Hogares afectados que sobredeclaran	1,21%	0,82%
Propensión a la ocultación		
Declaran propensión a ocultar	7,09%	5,22%
<0 <i>propensión a sobredeclarar</i>	<i>5,57%</i>	<i>3,04%</i>
Ocultación media de los que ocultan	2.077	1.745
<i>Sobredeclaración media de los que sobredeclaran</i>	<i>-1.374</i>	<i>-941</i>
No Declara propensión a ocultar	0,00%	0,00%
<i>propensión a sobredeclarar</i>	<i>0,00%</i>	<i>0,00%</i>
Ocultación media de los que ocultan	.	.
<i>Sobredeclaración media de los que sobredeclaran</i>	<i>.</i>	<i>.</i>
Declaran propensión a ocultar	0,54%	0,48%
>0 <i>propensión a sobredeclarar</i>	<i>0,97%</i>	<i>0,77%</i>
Ocultación media de los que ocultan	1.210	1.038
<i>Sobredeclaración media de los que sobredeclaran</i>	<i>-2.253</i>	<i>-1.468</i>

Fuente: Elaboración propia, a partir de las ECV y la CNE (base 2010).

Efecto sobre la Renta mixta neta

El cuadro 4.4 recoge el análisis de la principal fuente de renta de los trabajadores por cuenta propia (empresarios, autónomos y profesionales libres), la renta mixta.

En este caso, la nueva metodología de la base 2013 ha permitido reducir la ocultación inicialmente observada en la base 2004, aproximadamente a la mitad (50,62% en 2008 y 41,44% en 2011), habiéndose corregido más de las dos terceras partes del porcentaje de hogares que de entre el total están afectados por este tipo de renta (aproximadamente el 30%).

La corrección ha recaído sobre todos los tipos de afectados, hayan declarado positiva, negativamente o hayan ocultado totalmente sus rentas. Sin embargo, los afectados que no declararon este tipo de renta, ocultaron sus rentas porque eran bajas (corrección media de unos 7.000€) o incluso "perdonaron" ciertas pérdidas igualmente bajas (corrección media de unos -3.000€).

Se observa que la intensidad media de ocultación es alta (alrededor de los 23.000€) cuando se declararon cifras de ingresos negativos, bastante mayor que cuando se declararon ingresos positivos (alrededor de los 12.000€-13.000€).

Cuadro 4.4
EFFECTO SOBRE LA RENTA MIXTA NETA

B_3n	2008	2011
CNE (Residentes)	96.589.000.000	86.437.000.000
ECV (base 2004)	37.965.307.827	43.172.335.963
Ocultación	58.623.692.173	43.264.664.037
% Ocultación con respecto de CNE	60,69%	50,05%
ECV (base 2013)	67.638.559.753	61.102.168.874
Ocultación aflorada	36.128.117.956	30.162.546.955
<i>Sobredeclaración aflorada</i>	<i>-6.454.866.029</i>	<i>-12.232.714.045</i>
Corrección neta	29.673.251.926	17.929.832.910
% Corrección en esta variable / Su ocultación (b2004)	50,62%	41,44%
Hogares de la muestra (con mayores de 16 años)	13.360	12.714
% Hogares afectados de la muestra	30,50%	28,91%
% de Hogares afectados corregidos	73,01%	68,90%
% Hogares afectados que ocultan	60,10%	56,14%
% Hogares afectados que sobredeclaran	12,91%	12,76%
Propensión a la ocultación		
Declaran propensión a ocultar	85,81%	83,94%
<0 <i>propensión a sobredeclarar</i>	9,90%	12,05%
Ocultación media de los que ocultan	23.815	23.516
<i>Sobredeclaración media de los que sobredeclaran</i>	<i>-5.541</i>	<i>-1.667</i>
No Declara propensión a ocultar	8,90%	7,85%
<i>propensión a sobredeclarar</i>	0,64%	0,77%
Ocultación media de los que ocultan	7.875	7.378
<i>Sobredeclaración media de los que sobredeclaran</i>	<i>-3.327</i>	<i>-3.083</i>
Declaran propensión a ocultar	45,69%	41,86%
>0 <i>propensión a sobredeclarar</i>	15,40%	14,11%
Ocultación media de los que ocultan	13.328	12.229
<i>Sobredeclaración media de los que sobredeclaran</i>	<i>-10.377</i>	<i>-23.168</i>

Fuente: Elaboración propia, a partir de las ECV y la CNE (base 2010).

Efecto sobre los Sueldos y Salarios

El cuadro 4.5 muestra los resultados del análisis de las rentas de tipo salarial que perciben los asalariados (sin contar las cotizaciones sociales a cargo de los empleadores), que es la fuente de ingresos principal de los trabajadores por cuenta ajena.

Partiendo de una ocultación de en torno al 20%-25% en la base 2004 (20,73% en 2008 y 25,98% en 2011), esta partida ha visto reducida dicha ocultación en torno a un 40% en la base 2013 (41,80% en 2008 y 42,94% en 2011). En este caso, los hogares afectados por este tipo de renta suponen en torno a las dos terceras partes del total de ellos, viéndose corregidos en algún sentido la práctica totalidad de los mismos (más del 99% en ambos años).

Con respecto de las intensidades medias anuales de sus ocultaciones, son algo menores a las que presentaban los trabajadores por cuenta propia que declaraban cifras de ingresos positivas (en torno a los 9.000€ de media en ambos años); pero sustancialmente menores que las que presentaban los trabajadores por cuenta propia que declaraban pérdidas (cifras de ingresos negativas).

Cuadro 4.5
EFFECTO SOBRE LOS SUELDOS Y SALARIOS

D_11_r	2008	2011
CNE (Residentes)	442.319.000.000	420.917.000.000
ECV (base 2004)	350.613.064.932	311.559.338.141
Ocultación	91.705.935.068	109.357.661.859
% Ocultación con respecto de CNE	20,73%	25,98%
ECV (base 2013)	388.944.188.599	358.519.538.471
Ocultación aflorada	67.607.860.664	69.444.759.940
<i>Sobredeclaración aflorada</i>	<i>-29.276.736.997</i>	<i>-22.484.559.611</i>
Corrección neta	38.331.123.667	46.960.200.330
% Corrección en esta variable / Su ocultación (b2004)	41,80%	42,94%
Hogares de la muestra (con mayores de 16 años)	13.360	12.714
% Hogares afectados de la muestra	67,85%	63,31%
% de Hogares afectados corregidos	99,68%	99,55%
% Hogares afectados que ocultan	62,71%	67,00%
<i>% Hogares afectados que sobredeclaran</i>	<i>36,97%</i>	<i>32,55%</i>
Propensión a la ocultación		
No Declara propensión a ocultar	9,33%	11,58%
<i>propensión a sobredeclarar</i>	<i>0,00%</i>	<i>0,00%</i>
Ocultación media de los que ocultan	10.934	8.527
<i>Sobredeclaración media de los que sobredeclaran</i>		
Declaran propensión a ocultar	60,80%	64,29%
>0 <i>propensión a sobredeclarar</i>	<i>38,86%</i>	<i>35,22%</i>
Ocultación media de los que ocultan	9.139	8.993
<i>Sobredeclaración media de los que sobredeclaran</i>	<i>-6.068</i>	<i>-5.327</i>

Fuente: Elaboración propia, a partir de las ECV y la CNE (base 2010).

Sin embargo, hay aproximadamente un 10% de hogares de los que no declararon este tipo de renta que debieron hacerlo (9,33% y 11,58% en 2008 y 2011, respectivamente), siendo su nivel medio de ocultación significativo (10.934€ en 2008 y 8.527€ en 2011). En este caso, tanto la propensión como la intensidad media de ocultación son un poco mayores que en el caso de las rentas mixtas o por cuenta propia; posiblemente, porque en estos casos, se declararan pérdidas.

Efecto sobre las Cotizaciones sociales a cargo del empleador

El cuadro 4.6 muestra los resultados del análisis de la parte de las rentas de tipo salarial de los trabajadores por cuenta ajena que cotizan por ellos directamente sus empleadores.

Cabe destacar que alrededor de un 20% de los hogares que no declararon este tipo de renta la ocultaron totalmente. Partiendo de una ocultación de entre el 28%-30% en los años 2008 y 2011 según las ECV base 2004, la utilización de los registros administrativos ha conseguido reducir casi la mitad de ésta (45,97% para 2008 y 51,56% para 2011), llevándose a cabo estas correcciones sobre la totalidad de los hogares afectados por este tipo de rentas.

Se observa que la cantidad de hogares corregidos al alza por ocultación (53,48% y 59,02% en 2008 y 2011, respectivamente) es bastante parecida a la corregida a la baja por sobredeclaración (46,5% y 40,98%, respectivamente). No obstante, con respecto a las intensidades de las correcciones, las mayores correcciones se dan para los que no declararon este tipo de renta (7.265€ y

6.837€ en 2008 y 2011, respectivamente), siendo de menor importancia para los que las declararon, con correcciones medias superiores al alza por ocultación de en torno a los 3.200€ - 3.600€ y correcciones a la baja por sobredeclaración de en torno a los 1.900€ - 2.000€.

Cuadro 4.6
EFFECTO SOBRE LAS COTIZACIONES SOCIALES A CARGO DEL EMPLEADOR

D_12_r	2008	2011
CNE (Residentes)	118.198.000.000	110.959.000.000
ECV (base 2004)	84.146.232.214	74.491.179.414
Ocultación	34.051.767.786	36.467.820.586
% Ocultación con respecto de CNE	28,81%	32,87%
ECV (base 2013)	99.798.470.083	93.293.130.231
Ocultación aflorada	27.311.591.451	28.477.000.593
<i>Sobredeclaración aflorada</i>	<i>-11.659.353.582</i>	<i>-9.675.049.776</i>
Corrección neta	15.652.237.869	18.801.950.817
% Corrección en esta variable / Su ocultación (b2004)	45,97%	51,56%
Hogares de la muestra (con mayores de 16 años)	13.360	12.714
% Hogares afectados de la muestra	66,06%	62,07%
% de Hogares afectados corregidos	100,00%	100,00%
% Hogares afectados que ocultan	53,48%	59,02%
% Hogares afectados que sobredeclaran	46,52%	40,98%
Propensión a la ocultación		
No Declara propensión a ocultar	20,54%	21,49%
0 <i>propensión a sobredeclarar</i>	0,00%	0,00%
0 Ocultación media de los que ocultan	7.265	6.837
0 <i>Sobredeclaración media de los que sobredeclaran</i>		
Declaran propensión a ocultar	46,35%	50,78%
>0 <i>propensión a sobredeclarar</i>	53,65%	49,22%
0 Ocultación media de los que ocultan	3.625	3.246
0 <i>Sobredeclaración media de los que sobredeclaran</i>	-2.017	-1.907

Fuente: Elaboración propia, a partir de las ECV y la CNE (base 2010).

Efecto sobre los Intereses, Dividendos y Ganancias de Inversiones de Capital recibidos netos

En el cuadro 4.7 se muestran las Rentas procedentes del Saldo (Recursos-Empleos) resultante de la partida Intereses, Dividendos y Ganancias de Inversiones de Capital.

Como puede observarse en este cuadro, este tipo de rentas han sido sustancialmente incrementadas en 2011 mediante la corrección impuesta por la nueva metodología, rebajando el monto de su ocultación agregada en torno al 75% de la ocultación calculada para la base 2004 (que era muy considerable, del 81,99%).

Sin embargo, se observa que se ha producido un exceso de corrección en 2008, aflorando 19.190 millones de euros, cuando CNE recoge para esta partida un total agregado nacional de 11.701 millones de euros. Esta corrección debe ser, obviamente, puesta en cuestión.

El 100% de los hogares afectados fue corregido; su gran mayoría, al alza.

De los que no declararon este tipo de renta en la base 2004, aproximadamente, la mitad estaban ocultándolas, habiendo sido corregidos con una intensidad media de ocultación en torno a los 1.500 euros (1.660€ en 2008 y 1.474€ en 2011).

Cuadro 4.7
**EFFECTO SOBRE LOS INTERESES, DIVIDENDOS Y GANANCIAS DE
INVERSIONES DE CAPITAL RECIBIDOS NETOS**

D_4142_re	2008	2011
CNE (Residentes)	11.701.000.000	22.503.000.000
ECV (base 2004)	4.894.738.916	4.053.029.763
Ocultación	6.806.261.084	18.449.970.237
% Ocultación con respecto de CNE	58,17%	81,99%
ECV (base 2013)	22.361.623.358	17.930.409.696
Ocultación aflorada	19.190.024.534	15.235.880.752
<i>Sobredeclaración aflorada</i>	<i>-1.723.140.092</i>	<i>-1.358.500.818</i>
Corrección neta	17.466.884.442	13.877.379.933
% Corrección en esta variable / Su ocultación (b2004)	256,63%	75,22%
Hogares de la muestra (con mayores de 16 años)	13.360	12.714
% Hogares afectados de la muestra	64,84%	56,50%
% de Hogares afectados corregidos	100,00%	100,00%
% Hogares afectados que ocultan	83,79%	89,06%
% Hogares afectados que sobredeclaran	16,21%	10,94%
Propensión a la ocultación		
No Declara propensión a ocultar	50,14%	46,94%
0 <i>propensión a sobredeclarar</i>	0,00%	0,00%
0 Ocultación media de los que ocultan	1.660	1.474
0 <i>Sobredeclaración media de los que sobredeclaran</i>		
Declaran propensión a ocultar	64,34%	65,71%
>0 <i>propensión a sobredeclarar</i>	35,66%	34,29%
0 Ocultación media de los que ocultan	3.248	2.837
0 <i>Sobredeclaración media de los que sobredeclaran</i>	<i>-1.040</i>	<i>-1.309</i>

Fuente: Elaboración propia, a partir de las ECV y la CNE (base 2010).

Como consecuencia, aproximadamente las dos terceras partes de los hogares de la muestra (64,84% en 2008 y 56,50% en 2011) están afectadas por este tipo de renta; y han sido corregidos prácticamente todos, más de un 83% de ellos al alza, por motivo de ocultación (83,79% en 2008 y 89,06% en 2011).

En caso de haber declarado rentas positivas en la base de 2004, la intensidad media de ocultación fue prácticamente la doble que cuando no se declararon éstas (3.248€ frente 1.660€ en 2008, y 2.837€ frente a 1.474€ en 2011).

Efecto sobre el Saldo de rentas primarias netas

Como consecuencia de todo lo anterior, en el cuadro 4.8 se observa que la nueva metodología ha conseguido reducir algo más de la mitad la ocultación inicialmente existente en esta partida. Así, del 24,85% de ocultación observado en 2008, se ha corregido un 56,98%, y del 28,01% ocultado en 2011, se ha aflorado un 51,05%, afectando estas correcciones a casi todas las familias (98,38% en 2008 y 97,58% en 2011).

Si bien el número de hogares que no declaraban este tipo de renta y han sido corregidos es sustancialmente importante (47,97% y 43,76% de ellos en 2008 y 2011, respectivamente), sobre todo porque en ella se integran partidas mal recogidas en las ECV como los interés de capital, ganancias por inversiones o las cotizaciones sociales a cargo de los empleadores, la corrección en estos casos ha sido muy moderada (3.918€ y 5.408€ de media en 2008 y 2011, respectivamente). En estos casos, la propensión a sobredeclarar en los dos años estudiados fue casi nula (0,24% y 0,54%), con correcciones a la baja de medias 3.424€ y 2.684€.

Cuadro 4.8
EFFECTO SOBRE EL SALDO DE RENTAS PRIMARIAS NETAS

B_5n	2008	2011
CNE (Residentes)	725.636.000.000	698.849.000.000
ECV (base 2004)	545.347.482.806	503.117.688.454
Ocultación	180.288.517.194	195.731.311.546
% Ocultación con respecto de CNE	24,85%	28,01%
ECV (base 2013)	647.552.150.798	603.034.170.166
Ocultación aflorada	140.985.596.987	136.808.564.345
<i>Sobredeclaración aflorada</i>	<i>-38.780.928.995</i>	<i>-36.892.082.633</i>
Corrección neta	102.204.667.992	99.916.481.712
% Corrección en esta variable / Su ocultación (b2004)	56,69%	51,05%
Hogares de la muestra (con mayores de 16 años)	13.360	12.714
% Hogares afectados de la muestra	98,38%	97,58%
% de Hogares afectados corregidos	93,44%	90,29%
% Hogares afectados que ocultan	66,44%	67,61%
% Hogares afectados que sobredeclaran	27,00%	22,67%
Propensión a la ocultación		
Declaran propensión a ocultar	82,93%	80,60%
<0 <i>propensión a sobredeclarar</i>	12,20%	13,43%
0 Ocultación media de los que ocultan	33.246	19.117
0 <i>Sobredeclaración media de los que sobredeclaran</i>	-3.913	-3.308
No Declara propensión a ocultar	47,97%	43,76%
0 <i>propensión a sobredeclarar</i>	0,24%	0,54%
0 Ocultación media de los que ocultan	3.918	5.408
0 <i>Sobredeclaración media de los que sobredeclaran</i>	-3.424	-2.684
Declaran propensión a ocultar	65,76%	66,91%
>0 <i>propensión a sobredeclarar</i>	27,56%	23,16%
0 Ocultación media de los que ocultan	13.018	12.038
0 <i>Sobredeclaración media de los que sobredeclaran</i>	-7.692	-8.270

Fuente: Elaboración propia, a partir de las ECV y la CNE (base 2010).

Más fuerte ha sido la propensión y la corrección de los que declaraban cierta cantidad positiva en estas rentas, habiéndose corregido aproximadamente las dos terceras partes de estos (65,76% en 2008 y 66,91% en 2011) con afloramientos medios de 13.018€ y 12.038€ en 2008 y 2011, respectivamente. Pero también, en estos casos, se apreció cierta propensión a la sobredeclaración próxima a la cuarta parte de los hogares, con unas sobredeclaraciones medias de 7.692€ en 2008 y 8.270€ en 2011.

Pero, sin duda, las mayores correcciones se dieron para los hogares que declararon rentas negativas. De éstos, más del 80% ocultó rentas de forma reseñable. Así, el 82,93% de estos hogares ocultaron en promedio rentas del orden de los 33.246€ en 2008; y en 2011, fue el 80,60% de los hogares el que lo

hizo con una intensidad de ocultación media de 19.117€. Por el contrario, en estos casos, la propensión a sobredeclarar no llegó al 15% en ninguno de los dos años, siendo su sobredeclaración media similar a las de los hogares que no declararon este tipo de renta (alrededor de los 3.000€-4.000€).

Pasemos ahora a ver el efecto de la nueva metodología empleada en la base 2013 para corregir las partidas componentes de la Cuenta de Distribución Secundaria de la Renta.

Efecto sobre Impuestos corrientes + Cotizaciones sociales netas (empleos-recursos)

Considerando de forma agregada todos los impuestos, cuyas principales componentes son el Impuesto sobre la Renta (que la CNE recoge en la partida D_51_e) y el Impuesto sobre el Patrimonio (que la CNE incluye junto a otros en la partida D_59_e), junto con las Cotizaciones sociales⁶⁹ (que la CNE recoge en la partida D_61_e), el cuadro 4.9 nos muestra señales de las dificultades de recogida de esta partida.

Con la metodología de la base 2004, la agregación de estas partidas presenta unas ocultaciones del orden del 38,08% en 2008 y 40,47% en 2011, de las que la base 2013 solo consigue corregir un 39,81% de esta ocultación en 2008 y un 47,39% en 2011. En este sentido también las altas propensiones a ocultar en todas las situaciones (independientemente de que los hogares declaren valores positivos, negativos o simplemente no declaren para este tipo de renta) nos muestran indicios de una presumiblemente elevada falta de respuesta en esta partida.

Cuando los hogares declaraban negativamente en estas partidas de forma conjunta, en más del 88% de los casos se corrigió su declaración al alza con cantidades medias moderadas de unos 2.251€ en 2008 y 1.968€ en 2011, corrigiendo declaraciones a la baja con menor intensidad (en torno a los 650€-750€) en aproximadamente sólo el 10% de los casos.

La menor de las correcciones al alza por ocultación se realiza cuando no se declara este tipo de rentas (en torno a los 976€ en 2008 y 1.410€ en 2011), si bien, pese a su leve cuantía, afecta a más de la mitad de los casos (56,46% en 2008 y 54,22% en 2011).

Los mayores niveles de ocultación se dan para los que declaran cifras positivas para esta partida (6.757€ en 2008 y 6.840€ en 2011, de media), afectando aproximadamente al 60% de los casos en esta situación. Sin embargo, aunque menores, en este caso son también importantes y muy numerosas las correcciones a la baja por sobredeclaración, afectando aproximadamente al 40% de los casos y con unas correcciones medias a la baja de 2.938€ en 2008 y 2.728€ en 2011.

⁶⁹ En la ECV no se recogen los recursos de la partida *Cotizaciones sociales netas* de CNE.

Cuadro 4.9
EFECTO SOBRE IMPUESTOS CORRIENTES +
COTIZACIONES SOCIALES (EMPLEOS-RECURSOS)

D_5e_61_er	2008	2011
CNE (Residentes)	248.043.000.000	236.048.000.000
ECV (base 2004)	153.594.727.727	140.526.278.544
Ocultación	94.448.272.273	95.521.721.456
% Ocultación con respecto de CNE	38,08%	40,47%
ECV (base 2013)	191.195.022.259	185.796.944.917
Ocultación aflorada	56.539.818.043	61.069.000.060
<i>Sobredeclaración aflorada</i>	<i>-18.939.523.511</i>	<i>-15.798.333.687</i>
Corrección neta	37.600.294.533	45.270.666.373
% Corrección en esta variable / Su ocultación (b2004)	39,81%	47,39%
Hogares de la muestra (con mayores de 16 años)	13.360	12.714
% Hogares afectados de la muestra	95,89%	93,71%
% de Hogares afectados corregidos	99,83%	99,54%
% Hogares afectados que ocultan	61,77%	65,83%
% Hogares afectados que sobredeclaran	38,06%	33,71%
Propensión a la ocultación		
Declaran propensión a ocultar	88,05%	88,78%
<0 <i>propensión a sobredeclarar</i>	<i>10,32%</i>	<i>8,05%</i>
0 Ocultación media de los que ocultan	2.251	1.968
0 <i>Sobredeclaración media de los que sobredeclaran</i>	<i>-642</i>	<i>-766</i>
No Declara propensión a ocultar	56,46%	54,22%
0 <i>propensión a sobredeclarar</i>	<i>7,92%</i>	<i>7,63%</i>
0 Ocultación media de los que ocultan	976	1.410
0 <i>Sobredeclaración media de los que sobredeclaran</i>	<i>-408</i>	<i>-220</i>
Declaran propensión a ocultar	57,35%	60,84%
>0 <i>propensión a sobredeclarar</i>	<i>42,58%</i>	<i>38,88%</i>
0 Ocultación media de los que ocultan	6.757	6.840
0 <i>Sobredeclaración media de los que sobredeclaran</i>	<i>-2.938</i>	<i>-2.728</i>

Fuente: Elaboración propia, a partir de las ECV y la CNE (base 2010).

Efecto sobre las Prestaciones Sociales distintas de las transferencias sociales en especie (recursos)

Con respecto de las Prestaciones Sociales distintas de las Transferencias en especie, el cuadro 4.10 muestra sus resultados. Partiendo de unas ocultaciones relativas con respecto de lo marcado por CNE del 26,34% en 2008 y del 25,84% en 2011 registradas por la base 2004, las correcciones que permiten hacer la utilización de registros administrativos consiguen reducir éstas algo menos de la tercera parte (un 31,46% en 2008 y un 29,18% en 2011).

Aproximadamente un 70% de los hogares se ven afectados por este tipo de renta, donde se ubican las prestaciones por desempleo, jubilación, supervivencia, ayudas de estudio, etc. (un 69,21% en 2008 y un 72,90% en 2011). En consecuencia, estas cifras apuntan a que aproximadamente un 30% de los hogares españoles son contribuyentes netos que no perciben este tipo de prestaciones.

Casi la totalidad de los hogares afectados requirieron ser corregidos (un 92,58% de ellos en 2008 y un 94,42% de ellos en 2011). Y, como en casi todas las partidas anteriores que venimos explicando, son mayoría las correcciones por

ocultación con respecto a las realizadas por sobredeclaración -en este caso, en una relación aproximada de dos a uno-.

Con relación a las cuantías de las correcciones, cuando se declaraban rentas de este tipo, fueron casi tan importantes las correcciones por ocultación como las correcciones por sobredeclaración, con correcciones respectivas medias en torno a las 3.324€ al alza y 3.504€ a la baja en 2008, y de 3.520€ al alza y 4.027€ a la baja en 2011.

Cuadro 4.10
EFFECTO SOBRE PRESTACIONES SOCIALES DISTINTAS DE LAS TRANSFERENCIAS SOCIALES EN ESPECIE (RECURSOS)

D_62_r	2008	2011
CNE (Residentes)	154.572.000.000	183.845.000.000
ECV (base 2004)	113.851.048.059	136.341.174.588
Ocultación	40.720.951.941	47.503.825.412
% Ocultación con respecto de CNE	26,34%	25,84%
ECV (base 2013)	126.662.474.249	150.202.031.879
Ocultación aflorada	24.468.935.620	29.164.062.499
<i>Sobredeclaración aflorada</i>	<i>-11.657.509.431</i>	<i>-15.303.205.208</i>
Corrección neta	12.811.426.190	13.860.857.291
% Corrección en esta variable / Su ocultación (b2004)	31,46%	29,18%
Hogares de la muestra (con mayores de 16 años)	13.360	12.714
% Hogares afectados de la muestra	69,21%	72,90%
% de Hogares afectados corregidos	92,58%	94,42%
% Hogares afectados que ocultan	62,75%	64,33%
% Hogares afectados que sobredeclaran	29,83%	30,09%
Propensión a la ocultación		
No Declara propensión a ocultar	16,59%	21,78%
0 <i>propensión a sobredeclarar</i>	0,00%	0,00%
0 Ocultación media de los que ocultan	3.794	4.176
0 <i>Sobredeclaración media de los que sobredeclaran</i>		
Declaran propensión a ocultar	59,14%	60,22%
>0 <i>propensión a sobredeclarar</i>	32,72%	33,56%
0 Ocultación media de los que ocultan	3.324	3.520
0 <i>Sobredeclaración media de los que sobredeclaran</i>	<i>-3.504</i>	<i>-4.027</i>

Fuente: Elaboración propia, a partir de las ECV y la CNE (base 2010).

De órdenes parecidos fueron las correcciones cuando no se declararon este tipo de renta (3.794€ en 2008 y 4.176€ en 2011), si bien sólo se dieron al alza por ocultación y sólo en un 16,59% y 21,78% de los casos posibles, en los dos años estudiados.

Efecto sobre Otras transferencias corrientes

Por último, con respecto de las partidas "Otras transferencias corrientes", las más débilmente recogidas en la ECV y, como ahora veremos, corregidas, los cuadros 4.11 y 4.12 resumen sus principales resultados, el primero para las transferencias recibidas de otros hogares (recursos), y el segundo para las transferencias enviadas a otros hogares (empleos).

Cuadro 4.11
EFFECTO SOBRE OTRAS TRANSFERENCIAS CORRIENTES (RECURSOS)

D_7_r	2008	2011
CNE (Residentes)	51.565.000.000	47.374.000.000
ECV (base 2004)	2.273.759.577	1.984.741.972
Ocultación	49.291.240.423	45.389.258.028
% Ocultación con respecto de CNE	95,59%	95,81%
ECV (base 2013)	2.263.734.161	1.978.490.086
Ocultación aflorada	1.531.468	802.509
<i>Sobredeclaración aflorada</i>	-11.556.883	-7.054.394
Corrección neta	-10.025.415	-6.251.886
% Corrección en esta variable / Su ocultación (b2004)	-0,02%	-0,01%
Hogares de la muestra (con mayores de 16 años)	13.360	12.714
% Hogares afectados de la muestra	3,37%	2,92%
% de Hogares afectados corregidos	3,11%	3,77%
% Hogares afectados que ocultan	0,67%	2,16%
% Hogares afectados que sobredeclaran	2,44%	1,62%
Propensión a la ocultación		
No Declara propensión a ocultar	0,00%	0,00%
0 <i>propensión a sobredeclarar</i>	0,00%	0,00%
0 Ocultación media de los que ocultan	.	.
0 <i>Sobredeclaración media de los que sobredeclaran</i>		
Declaran propensión a ocultar	0,67%	2,16%
>0 <i>propensión a sobredeclarar</i>	2,44%	1,62%
0 Ocultación media de los que ocultan	281	127
0 <i>Sobredeclaración media de los que sobredeclaran</i>	-619	-535

Fuente: Elaboración propia, a partir de las ECV y la CNE (base 2010).

Como puede observarse en el cuadro 4.11 relativo a los recursos correspondientes a esta partida (D_7_r - Otras transferencias corrientes recibidas por los hogares), en términos relativos, las deficiencias son importantes. Las ocultaciones presentadas en la base 2004 son la casi totalidad de la partida (el 95,59% en 2008 y el 95,81% en 2011).

Pero, desgraciadamente, la utilización de registros administrativos en la base 2013, prácticamente, no mejora en nada la situación. Las correcciones realizadas son puramente testimoniales. De hecho, no se descubre ningún nuevo hogar que no declarara este tipo de renta y haya sido corregido. Y en los pocos hogares que están afectados por este tipo de transferencias por el hecho de que declararon alguna cantidad en ella (un 3,37% en 2008 y un 2,92% en 2011) se detectan más sobredeclaraciones que ocultaciones, produciéndose agregadamente un incremento neto de la ocultación con respecto de lo que marca la CNE (incremento de un 0,02% en 2008 e incremento de un 0,01% en 2011).

Y algo parecido ocurre con los empleos (D_7_e - Otras transferencias corrientes enviadas por los hogares), como podemos observar en el cuadro 4.12. Partiendo de unas ocultaciones en base 2004 del 92,31% en 2008 y del 93,64% en 2011, el empleo de los registros administrativos solo permite descubrir un 0,10% en 2008 y un 0,08% en 2011 de hogares que no declaran este tipo de renta y debían haberlo hecho, con una ocultación media de 4.415€ en 2008 y 1.780€ en 2011.

Cuadro 4.12
EFECTO SOBRE OTRAS TRANSFERENCIAS CORRIENTES (EMPLEOS)

D_7_e	2008	2011
CNE (Residentes)	54.594.000.000	54.811.000.000
ECV (base 2004)	4.200.439.344	3.486.852.198
Ocultación	50.393.560.656	51.324.147.802
% Ocultación con respecto de CNE	92,31%	93,64%
ECV (base 2013)	4.355.667.692	3.571.811.227
Ocultación aflorada	205.146.009	122.583.139
<i>Sobredeclaración aflorada</i>	<i>-49.917.661</i>	<i>-37.624.110</i>
Corrección neta	155.228.348	84.959.029
% Corrección en esta variable / Su ocultación (b2004)	0,31%	0,17%
Hogares de la muestra (con mayores de 16 años)	13.360	12.714
% Hogares afectados de la muestra	6,19%	5,49%
% de Hogares afectados corregidos	10,04%	11,89%
% Hogares afectados que ocultan	6,65%	8,02%
% Hogares afectados que sobredeclaran	3,39%	3,87%
Propensión a la ocultación		
No Declara propensión a ocultar	0,10%	0,08%
0 <i>propensión a sobredeclorar</i>	0,00%	0,00%
0 Ocultación media de los que ocultan	4.415	1.780
0 <i>Sobredeclaración media de los que sobredeclaran</i>		
Declaran propensión a ocultar	5,28%	6,69%
>0 <i>propensión a sobredeclorar</i>	3,44%	3,92%
0 Ocultación media de los que ocultan	1.733	1.130
0 <i>Sobredeclaración media de los que sobredeclaran</i>	<i>-976</i>	<i>-814</i>

Fuente: Elaboración propia, a partir de las ECV y la CNE (base 2010).

En resumen, sólo un 6,19% de los hogares en 2008 y un 5,49% en 2011 se vieron afectados por este tipo de renta, y de ellos, sólo se corrigieron aproximadamente un 10%, reduciéndose la ocultación de la base 2004 en un 0,31% en 2008 y en un 0,17% en 2011; correcciones puramente testimoniales como veníamos diciendo.

Es importante destacar que, aunque en términos de participación en la renta disponible a nivel agregado, estas partidas no son tan importantes como la mayoría de las restantes, y aunque a nivel de Contabilidad Nacional a veces se obvia su importancia con la excusa de que los empleos y los recursos recogidos en estas partidas se compensan y prácticamente se anulan en el territorio nacional, la verdad es que, desde el punto de vista microeconómico, estas partidas son ciertamente importantes. Debido a la asimetría con que redistribuyen las rentas entre los hogares que las reciben o las envían, pueden movilizar sustancialmente a los hogares sobre la escala de las rentas, tanto más cuanto mayor sea la cuantía de sus transferencias.

Por tanto, en nuestra opinión, es ésta una verdadera asignatura pendiente en estas encuestas que necesitamos superar para describir con realismo la distribución de los tamaños de las rentas de las personas y hogares. En cualquier caso, en lo que resta de este trabajo, trataremos de emplear la información disponible fusionada al respecto con el objetivo de mejorar esta partida.

Efecto sobre la Renta disponible neta

Siendo la corrección realizada sobre la Renta Disponible el resultado de las correcciones de los distintos tipos de renta que la componen, sus resultados reflejan en cierta forma el comportamiento "promedio" de las correcciones realizadas sobre cada una de sus partidas o componentes y que han sido presentadas en los cuadros 4.2 a 4.12.

Así pues, como resumen final, el cuadro 4.13 muestra cómo la nueva metodología ha conseguido reducir la ocultación que se presentaba en la base 2004 para la Renta Disponible neta, a aproximadamente la mitad (partiendo de ocultaciones del 19,89% en 2008 y del 22,16% en 2011, ha reducido estas cantidades en un 61,17% en 2011 y un 48,31% en 2008). La verdadera ocultación aflorada es mayor que estas correcciones en los dos años; pero además las cuantías agregadas de las correcciones a la baja para las sobredeclaraciones también detectadas son también considerables, aminorando el saldo final agregado de la corrección realizada.

La práctica totalidad de los hogares han sido corregidos en algún sentido (99,93% en 2008 y 99,82% en 2011), la gran mayoría de ellos, por ocultación. En concreto, en los dos años considerados, se han corregido por ocultación el 84,76% y el 87,10%, respectivamente, de los que declaran negativamente, el 70,68% y el 72,25%, respectivamente, de los que declaran negativamente, y el 70,59% y el 56,36%, respectivamente, de los que no declaran).

Se observa que tanto la incidencia del fenómeno (propensión a ocultar) como la intensidad del mismo (ocultación media de los que ocultan) es mucho mayor cuando se declaran rentas negativas (correcciones medias mayores de 32.000€) que cuando no cuando se declaran rentas o se declaran rentas positivas (correcciones medias no superiores a los 9.000€ aproximadamente en ambos casos).

Sin embargo, esta diferencia de comportamiento o tendencia no se aprecia cuando nos fijamos en las sobredeclaraciones corregidas, que presenta valores más parecidos (entre 3.300€ y 6.500€).

Cuadro 4.13
EFECTO SOBRE RENTA DISPONIBLE NETA

B_6n	2008	2011
CNE (Residentes)	628.767.000.000	639.038.000.000
ECV (base 2004)	503.677.123.370	497.430.474.272
Ocultación	125.089.876.630	141.607.525.728
% Ocultación con respecto de CNE	19,89%	22,16%
ECV (base 2013)	580.927.669.257	565.845.935.987
Ocultación aflorada	108.234.487.518	101.939.087.573
<i>Sobredeclaración aflorada</i>	<i>-30.983.941.632</i>	<i>-33.523.625.857</i>
Corrección neta	77.250.545.887	68.415.461.715
% Corrección en esta variable / Su ocultación (b2004)	61,76%	48,31%
Hogares de la muestra (con mayores de 16 años)	13.360	12.714
% Hogares afectados de la muestra	99,93%	99,82%
% de Hogares afectados corregidos	99,63%	99,41%
% Hogares afectados que ocultan	70,85%	72,38%
% Hogares afectados que sobredeclaran	28,79%	27,03%
Propensión a la ocultación		
Declaran propensión a ocultar	84,76%	87,10%
<0 <i>propensión a sobredeclarar</i>	15,24%	9,68%
0 Ocultación media de los que ocultan	33.036	32.117
0 <i>Sobredeclaración media de los que sobredeclaran</i>	-6.947	-3.318
No Declara propensión a ocultar	70,59%	56,36%
0 <i>propensión a sobredeclarar</i>	0,00%	1,82%
0 Ocultación media de los que ocultan	4.521	8.281
0 <i>Sobredeclaración media de los que sobredeclaran</i>		-53
Declaran propensión a ocultar	70,68%	72,25%
>0 <i>propensión a sobredeclarar</i>	28,95%	27,17%
0 Ocultación media de los que ocultan	9.005	7.975
0 <i>Sobredeclaración media de los que sobredeclaran</i>	-5.704	-6.410

Fuente: Elaboración propia, a partir de las ECV y la CNE (base 2010).

3.2.1. Algunas conclusiones

Efectivamente, la nueva metodología empleada en las ECV parece haber mejorado sustancialmente la calidad de la información sobre los ingresos que recoge. Y el análisis de los microdatos fusionados de ambas bases nos sugiere una futura vía de investigación del fenómeno de la subdeclaración en las encuestas, así como de la caracterización del fenómeno de la ocultación. De esta forma, vinculando las cantidades ocultas por los hogares con sus características podríamos determinar cuáles de éstas podrían ser de ayuda para predecir, tanto la propensión a subdeclarar de los individuos en los diferentes tipos de renta, como la intensidad con que se daría dicha subdeclaración u ocultación en el caso de que la persona realmente subdeclarara.

Sus resultados permitirían imputar racionalmente los restos de ocultación que trascienden a los registros administrativos y que la CNE detecta, propiciando distribuciones personales de rentas más coherentes con el marco contable que admitimos como referencia.

Sin embargo, en unos primeros análisis comparativos de los resultados de las dos bases de las ECV que hemos realizado en el transcurso del tiempo de

realización de esta tesis (Callealta y Romero, 2016), se mostraban algunos resultados un tanto llamativos, afortunadamente pocos, que merecen alguna investigación más para conseguir explicarlos y aislar sus efectos. Entre ellos, se decía:

Al comparar... las ocultaciones en rentas mixtas de las diferentes regiones, y observamos que las mínimas incidencias e intensidades de la ocultación se dan sistemáticamente en el País Vasco, donde además, las correcciones son predominantemente por motivo de sobredeclaración.

Inversamente y con relación a las rentas salariales, impuestos corrientes y cotizaciones sociales,... otra vez es en el País Vasco donde se realizan las correcciones más fuertes (mucho mayores en el caso de las rentas salariales), ahora por motivos de ocultación predominantemente.

Pareciera como si la información administrativa del País Vasco no estuviese fluyendo hacia las ECV como lo hace la del resto del territorio nacional. Y, efectivamente, el INE mantiene actualmente convenios de colaboración en materia estadística con la Agencia Estatal de Administración Tributaria, la Seguridad Social, la Hacienda Tributaria de Navarra y la Diputación Foral de Bizkaia. Es decir, que no hay información fiscal de todo el País Vasco, y mientras esto sea así, se estará introduciendo un sesgo en las correcciones que afecta a la distribución de las rentas de los hogares, en virtud de su residencia.

Y también en aquél estudio se decía:

... cuando se consideran las incidencias e intensidades de la ocultación de la renta disponible neta de los hogares... de los hogares formados por una única persona de menos de 30 años prácticamente nunca ocultan.

Quizás esto se deba a que sus rentas son relativamente bajas y no estén obligados a declarar. Pero, aún en este caso, los registros administrativos recogen información fiscal al respecto. Quizás se hayan imputado rentas en exceso a este colectivo en la base 2004. En cualquier caso, este punto requiere una mayor investigación para aclararlo.

En conclusión, y a pesar de algunos de estos posibles sesgos que la nueva metodología pudiera estar introduciendo, desde la perspectiva de los ingresos, la incorporación de los registros administrativos a las ECV parece haber sido muy positiva. Las comparaciones realizadas entre las ECV base 2013 y las ECV base 2004 a nivel agregado, nos permiten afirmar que la utilización de los registros administrativos ha permitido elevar los niveles de renta estimados para las personas y hogares, aproximándolos más a los que muestra la CNE. Esto ocurre no sólo para la renta disponible agregada de los hogares, sino también para las diversas partidas de renta que la integran, con la excepción de "Otras transferencias corrientes" que siguen constituyendo una verdadera asignatura pendiente para estas encuestas.

Pero el fenómeno de la subdeclaración u ocultación de rentas trasciende a la propia ECV y se aprecia en todas las estadísticas sobre ingresos, incluidos los propios registros administrativos. Y, a pesar de las mejoras constatadas en la nueva metodología empleada por las ECV base 2013, éstas siguen

subdeclarando los niveles de renta agregada que la CNE muestra; si bien en menor medida y de modo más homogéneo para las diferentes partidas de lo que lo hacía la anterior metodología (base 2004).

Además, el análisis ha desvelado claramente que la aplicación de esta nueva metodología no sólo ha descubierto las perseguidas ocultaciones de rentas en los hogares, sino que también ha desvelado una no despreciable sobredeclaración de rentas; lo que contraviene las expectativas de sobredeclaración imputables a los efectos del redondeo u imprecisiones por olvido lógicamente esperables en declaraciones del tipo como las que recoge las ECV, sino que presumiblemente estén también corrigiendo cierta sobreimputación de rentas realizada en la base 2004, ante la presencia de una masiva falta de respuesta en muchos casos.

Así las cosas, la nueva metodología supone un buen paso en la búsqueda de una mejor estimación de la distribución de las rentas. Pero una estimación más realista de la distribución de las rentas reales, no simplemente declaradas, de las personas y hogares exige seguir avanzando en superar las debilidades informativas que las encuestas aún presentan, en la investigación de la conducta de las personas que ocultan rentas cuando las deben declarar, así como en los posibles efectos que los procedimientos de imputación estén pudiendo imprimir sobre los microdatos generados. Lógicamente, este objetivo trasciende el objetivo primario y la extensión de esta Tesis.

Pero, en este sentido y ante la ausencia de otros microdatos públicos que permitan investigar más precisamente estos aspectos, trataremos en esta tesis de realizar una primera aplicación de esta posibilidad que nos brindan las dos bases publicadas para las ECV. Aplicando la misma metodología empleada para el estudio presentado a lo largo de este apartado nº 3 de este capítulo, dispondremos, a nivel de microdato y para los años de las olas 2008-2012, de las cantidades presumiblemente ocultas por las personas y hogares encuestados en cada tipo de ingreso, y que supondremos que son, mínimamente, las ocultaciones desveladas estrictamente positivas (las sobredeclaraciones, negativas, serían asimilables a fenómenos de olvido o sobreimputación). Asumidas estas ocultaciones mínimas, las emplearemos para tratar de profundizar en la objetivación del modelo de corrección propuesto en el apartado 2 y, particularmente, para estimar objetivamente la a priori indeterminada propensión a la ocultación, k . Presentaremos esta aplicación en el apartado 4.2 de este capítulo.

4. Implementación práctica del modelo propuesto para la estimación combinada de la DPR

De acuerdo con lo que hemos comentado en los apartados anteriores, la aplicación de la metodología propuesta se apoya en tres pilares:

- el modelo, ya propuesto en el apartado 2.2 de este capítulo;
- los indicadores de incongruencia calculados con la información complementaria allegada de las fuentes fusionadas en el capítulo 2 de este trabajo, empleando la metodología expuesta en el apartado 2.3;

- y la información complementaria sobre ocultación de rentas que permita, empleando la metodología propuesta en el apartado 2.4, estimar objetivamente para cada clase el parámetro "k" indeterminado para el modelo; y que, en nuestro caso, extraeremos de lo expuesto en el apartado 3.

Veamos, pues, a continuación cómo concretar la implementación práctica de estos dos últimos pilares.

4.1. Determinación de los niveles de incongruencia de las rentas declaradas de los hogares

En nuestro modelo intervienen dos tipos de indicadores de incongruencia:

- Los que se construyen específicamente para cada tipo de renta declarada que se pretende corregir (en el modelo notados por X), y que proveen básicamente el grado de corrección (la corrección es proporcional a estos).
- Los que se construyen como indicadores generales de ocultación de rentas agregadas en el hogar (en el modelo notados por Y), independientemente de su tipología, y que se propone emplear como moduladores de la constante de proporcionalidad (determinan la inclinación de la forma lineal del factor de proporcionalidad).

4.1.1. Indicadores de incongruencia específicos de cada partida

Entre las diferentes partidas de renta de los hogares, encontramos algunas de ellas que son derivadas directamente de otras más básicas y que serían las que requieren la atención a la hora de realizar las pertinentes correcciones por ocultación.

En los siguientes cuadros se presentan los indicadores que hemos construido para cada una de las partidas de rentas y sus distribuciones, divididos por su extensión de acuerdo con si son propias de la Cuenta de Asignación de la Renta Primaria (cuadros 4.14.a, b y c), o de la Cuenta de Distribución Secundaria de la Renta (cuadros 4.14.d, e y f).

En ellos se especifican, para cada partida de renta en la Cuenta de Rentas del Sector Hogares (CRSH) construida en nuestra Base de Datos de Hogares (BDH).

Así, en la columna "Variables BDH a Corregir/Calcular" se presentan los nombres de las variables en que se recogen dichas partidas en nuestra Base de Datos (BDH). Se muestran con fondo de color amarillo si necesitan corrección directa, mientras que el fondo verde nos indica si dichas partidas se corrigen indirectamente como consecuencia de un cálculo a partir de otras variables corregidas previamente. Las variables cuyos nombres se presentan en color naranja, requerirán ser corregidas en una segunda vuelta o tercera

vueltras, pues para ello necesitan disponer de ciertas variables que deben haber sido corregidas previamente. Tal es el caso de las Cotizaciones Sociales a cargo de los empleadores, que requieren conocimiento previo de los Sueldos y Salarios, por ejemplo.

Las relaciones que se espera que mantengan con otras variables que se emplearán para calcular los correspondientes Indicadores de Incongruencia para aquéllas, se presentan en la correspondiente columna así descrita en su cabecera. Un conector (+) entre el nombre de la variable declarada a corregir y la variable presumiblemente relacionada ordinalmente con ella, indica que se espera una relación positiva o directa entre las mismas; y, en este caso, la relación se marca en color azul. Por el contrario, un conector (-) entre el nombre de la variable declarada a corregir y la variable presumiblemente relacionada ordinalmente con ella indica que se espera una relación inversa o negativa entre las mismas; marcándose la relación, en este caso, en color verde. En fondo verdoso se marcan las fórmulas de cálculo de aquellas partidas en las que sus correspondientes variables no necesitan ser corregidas directamente, sino que se corregirán indirectamente a través de esta expresión.

Como información complementaria a la anterior columna, bajo la cabecera "Observaciones" se describen someramente las descripciones de las variables que intervienen en las relaciones ordinales que presumiblemente mantienen con la variable que recoge la partida correspondiente.

Cuadro 4.14.a

INDICADORES ESPECÍFICOS (CUENTA DE ASIGNACIÓN DE LA RENTA PRIMARIA)

Partidas de la CRSH en la BDH		Variables BDH a Corregir / Calcular	Relaciones entre Variables usadas para calcular los Indicadores de Congruencia	Observaciones	Indicadores de Incongruencia	Filtro considerado para la aplicación de la corrección
Excedente de Explotación	Producción relacionada con el Alquiler imputado Vivienda Principal (si el Hogar la ocupa en propiedad)	AIVPP_P	$AIVPP_P (+) P2_5_eff * Ind(P2_1_eff=2)$	P2_1_eff=2: Propiedad (incluye herencia y regalo) P2_5_eff: Valor actual de la vivienda principal regteneviv=1: Vivienda habitual en propiedad	B2_001	regteneviv=1
	Intereses pagados de préstamos hipotecarios	AIVPP_CI_IPPH	$AIVPP_CI_IPPH (+) IPPH_B_EFF * Ind(P2_1_eff=2)$	regteneviv=2: Vivienda habitual en propiedad, pero con hipoteca	B2_002	regteneviv=2
	Servicios de Seguros y Planes de Pensiones	AIVPP_CI_Seg	$AIVPP_CI_Seg (+) (P5_20b_eff + P5_20c_eff) * Ind(P2_1_eff=2)$	P5_20b_eff: Importe prima anual de Seguros de Vida No voluntarios P5_20c_eff: Importe pago único de Seguros de Vida No voluntarios	B2_013	regteneviv=2
			$AIVPP_CI_Seg (+) IPPH_B_EFF * Ind(P2_1_eff=2)$	IPPH_B_EFF: Intereses pagados por préstamos hipotecarios [EFF]	B2_003	
				Primera componente principal de B2_013 y B2_003	B2_000	
	Consumos Intermedios relacionados con el AIVPP	AIVPP_CI	$AIVPP_CI = AIVPP_CI_IPPH + AIVPP_CI_Seg$			
	Otros Impuestos netos sobre la Producción relacionados con el AIVPP	AIVPP_OIN	$AIVPP_OIN (+) P2_5_eff * Ind(P2_1_eff=2)$	P2_1_eff=2: Propiedad (incluye herencia y regalo) P2_5_eff: Valor actual de la vivienda principal	B2_004	regteneviv=1
	Excedente Bruto de Explotación	B_2b_Bdh	$B_2b_Bdh = AIVPP_P - AIVPP_CI - AIVPP_OIN$			
Renta mixta	Renta Mixta Neta	B_3n_Bdh	$B_3n_Bdh (+) RMIX_B_EFF + IMALQPRO_EFF$	RMIX_B_EFF: Ingresos brutos por cuenta propia (EFF) IMALQPRO_EFF: Ingresos por Alquiler de Propiedades, excluido terrenos (EFF)	B3_001	$B_3n_Bdh > 0$ or $(RMIX_B_EFF + IMALQPRO_EFF > 0)$
			$B_3n_Bdh (+) V_CUEYDEP_eff + sum(P3_5_i_eff, i=1-8)$	V_CUEYDEP_eff: Valor de cuentas y depósitos utilizables para realizar pagos P3_5_i_eff, (i=1-8): importe inicial Préstamos para hogares o empresas individuales no destinados a vivienda, como indicador de solvencia	B3_002	
				Primera componente principal de B3_001 y B3_002	B3_000	
	Consumo de Capital Fijo de la Renta Mixta	CCFRM	$CCFRM (+) sum(P4_108_i * (1.0208)^{(ano-P4_107_i) * min(1; (ano-P4_107_i)/20)})$	P4_108_i_eff, (i=1-6): Valor inicial de los negocios P4_107_i_eff, (i=1-6): año de comienzo de los negocios	B3_003	$B_3n_Bdh > 0$ or $(RMIX_B_EFF + IMALQPRO_EFF > 0)$
	Renta mixta bruta	B_3b_Bdh	$B_3b_Bdh = B_3n_Bdh + CCFRM$			
Remuneración de Asalarados	Sueldos y Salarios	D_11_r_Bdh	$D_11_r_Bdh (+) RMCA_B_EFF + RNMCA_B_EFF$	RMCA_B_EFF: Renta Bruta Monetaria por cuenta ajena (EFF) RNMCA_B_EFF: Renta Bruta No Monetaria por cuenta ajena (EFF)	D1_001	$D_11_r_Bdh > 0$ or $(RMCA_B_EFF + RNMCA_B_EFF > 0)$
	Colizaciones sociales efectivas a cargo de empleadores	D_121_r_Bdh	$D_121_r_Bdh (+) D_11_r_Bdh$		D1_002	$D_11_r_Bdh > 0$ or $(RMCA_B_EFF + RNMCA_B_EFF > 0)$
	Colizaciones sociales imputadas a cargo de empleadores	D_122_r_Bdh	$D_122_r_Bdh (+) D_11_r_Bdh / nmh$		D1_003	$D_11_r_Bdh > 0$ or $(RMCA_B_EFF + RNMCA_B_EFF > 0)$
	Colizaciones sociales a cargo del empleador	D_12_r_Bdh	$D_12_r_Bdh = D_121_r_Bdh + D_122_r_Bdh$			
	Remuneración de Asalarados	D_1_r_Bdh	$D_1_r_Bdh = D_11_r_Bdh + D_12_r_Bdh$			
Rentas de la propiedad	Intereses (recursos) + Rentas distribuidas de las sociedades (recursos)	D_4142_r_Bdh	$D_4142_r_Bdh (+) IMCAP_B_EFF$	IMCAP_B_EFF: Rentas Brutas del capital (EFF)	D4_001	$D_4142_r_Bdh > 0$ or $IMCAP_B_EFF > 0$ or $actfinanc_eff > 0$
			$D_4142_r_Bdh (+) actfinanc_eff$	actfinanc_eff: Valor Total de los ACTIVOS FINANCIEROS	D4_002	
				Primera componente principal de D4_001 y D4_002	D4_000	
	Intereses (empleos)	D_41_e_Bdh	$D_41_e_Bdh (+) deuotros_eff$	deuotros_eff: Valor de Deudas NO INMOBILIARIAS	D4_003	$D_41_e_Bdh > 0$ or $deuotros_eff > 0$
	Saldo (Rec-Emp) de Intereses, Dividendos y Ganancias de Inversiones de Capital	D_4142_re_Bdh	$D_4142_re_Bdh = D_4142_r_Bdh - D_41_e_Bdh$			
	Otras rentas de inversión	D_44_r_Bdh	$D_44_r_Bdh (+) all_eff + valor_eff + valseg_eff$	all_eff: Valor de FONDOS DE INVERSION valseg_eff: Valor de SEGUROS DE VIDA valor_eff: Valor de PLANES DE PENSIONES	D4_004	$D_44_r_Bdh > 0$ or $actfinanc_eff > 0$
	Rentas de la tierra (rec-emp)	D_45_re_Bdh	$D_45_re_Bdh (+) IMALQTER_EFF$	IMALQTER_EFF: Ingresos por Alquiler de Terrenos (EFF)	D4_006	$D_45_re > 0$ or $IMALQTER_EFF > 0$
Saldo de rentas primarias	Rentas de la propiedad (rec-emp)	D_4_re_Bdh	$D_4_re_Bdh = D_4142_re_Bdh + D_44_r_Bdh + D_45_re_Bdh$			
	Saldo de rentas primarias brutas	B_5b_Bdh	$B_5b_Bdh = B_2b3b_Bdh + D_1_r_Bdh + D_4_re_Bdh$			

Fuente: Elaboración propia a partir de la CNE, las ECV y las EFF.

Como consecuencia, cada una de estas relaciones será empleada para calcular los tres indicadores (uno de cada tipo, I_c , D_c y J_c , presentados, en el apartado 2.3), teniendo en cuenta el tipo de relación (positiva o negativa) entre las variables que los generarán, así como el signo de la ocultación ($O > 0$ u ocultación, $O < 0$ o sobredeclaración) para cuya corrección se emplearán, y que se calcularán según lo descrito ya en el mismo apartado. A partir de éstos, finalmente, se obtendrá el correspondiente Indicador de incongruencia para las declaraciones de los casos en la correspondiente partida, calculado como la primera componente principal de aquéllos, como allí se indicó y como expresión sintética común del nivel de incongruencia compartido por todos

ello. El nombre que daremos finalmente a dicho indicador se muestra en la correspondiente fila de la columna así encabezada.

Por último, se indica en cada caso, el filtro que se considerará para seleccionar los hogares que deben ser sometidos de forma efectiva al proceso de corrección, dependiendo de la casuística de percepción, o no, de los tipos de rentas consideradas.

En los dos siguientes cuadros se muestran, respectivamente, los estadísticos básicos obtenidos para las distribuciones de valores de los indicadores presentados en el cuadro 4.14.a, una vez que fueron calculados cuando partíamos de los datos de las ECV base 2004 (cuadro 4.14.b) y de los de las ECV base 2013 (cuadro 4.14.c). Ellos nos permiten analizar someramente sus distribuciones.

Cuadro 4.14.b

INDICADORES ESPECÍFICOS ASOCIADOS A LAS PARTIDAS DE LA CUENTA DE ASIGNACIÓN DE LA RENTA PRIMARIA, CALCULADOS PARA LA BASE 2004

Indicadores	Mínimo	Percentiles											Máximo
		1	5	10	25	50	75	90	95	99	99,5	99,9	
B2_001	-0,593	-0,593	-0,593	-0,578	-0,492	-0,327	0,042	0,592	1,121	2,726	3,779	8,135	127,474
B2_002	-0,715	-0,715	-0,715	-0,715	-0,715	-0,672	-0,672	-0,183	0,387	2,331	3,213	5,496	16,449
B2_013	-0,682	-0,682	-0,682	-0,667	-0,433	-0,090	-0,090	0,425	0,977	2,350	3,500	9,223	116,227
B2_003	-0,730	-0,730	-0,730	-0,718	-0,498	-0,265	-0,057	0,735	1,416	3,181	3,969	6,772	137,747
B2_000	-0,655	-0,655	-0,655	-0,638	-0,366	-0,223	0,059	0,813	1,424	3,078	4,142	8,180	102,387
B2_004	-0,934	-0,934	-0,934	-0,901	-0,499	-0,358	0,094	0,819	1,430	3,098	4,006	7,661	71,484
B2_005	-0,892	-0,892	-0,892	-0,866	-0,497	-0,337	0,082	0,751	1,404	3,079	3,973	7,133	86,328
B3_001	-0,244	-0,244	-0,244	-0,244	-0,244	-0,152	-0,152	-0,017	0,108	0,495	0,828	3,629	102,572
B3_002	-0,308	-0,308	-0,308	-0,308	-0,308	-0,275	-0,239	0,023	0,217	0,648	0,921	2,703	61,682
B3_000	-0,198	-0,198	-0,198	-0,198	-0,198	-0,175	-0,040	0,401	0,694	1,565	2,230	7,548	118,922
B3_003	-0,078	-0,078	-0,078	-0,078	-0,078	-0,050	-0,050	-0,050	-0,048	0,375	0,770	1,921	82,306
D1_001	-0,312	-0,312	-0,312	-0,312	-0,278	-0,176	-0,064	0,119	0,300	0,989	1,557	4,807	157,821
D1_002	-0,205	-0,205	-0,205	-0,205	-0,198	-0,142	-0,100	0,097	0,336	1,216	1,546	2,602	182,557
D1_003	-1,000	-1,000	-1,000	-1,000	-1,000	-0,978	0,048	1,074	1,617	2,523	2,780	3,310	3,878
D4_001	-0,096	-0,096	-0,096	-0,096	-0,096	-0,060	0,022	0,088	0,123	0,388	0,791	3,466	107,430
D4_002	-0,170	-0,170	-0,164	-0,157	-0,138	-0,078	0,018	0,106	0,173	0,747	1,513	10,045	128,423
D4_000	-0,153	-0,153	-0,149	-0,138	-0,114	-0,073	0,011	0,108	0,170	0,685	1,345	7,612	101,513
D4_003	-0,142	-0,142	-0,142	-0,142	-0,142	-0,084	-0,048	0,042	0,091	0,245	0,393	1,217	138,711
D4_004	-0,233	-0,233	-0,233	-0,218	-0,157	-0,087	0,031	0,178	0,318	1,067	1,672	7,184	90,856
D4_006	-0,431	-0,431	-0,431	-0,431	-0,431	-0,262	-0,262	-0,262	-0,262	-0,262	-0,262	0,020	9,587

Fuente: Elaboración propia.

Cuadro 4.14.c
INDICADORES ESPECÍFICOS ASOCIADOS A LAS PARTIDAS DE LA CUENTA DE
ASIGNACIÓN DE LA RENTA PRIMARIA, CALCULADOS PARA LA BASE 2013

Indicadores	Mínimo	Percentiles											Máximo
		1	5	10	25	50	75	90	95	99	99,5	99,9	
B2_001	-0,524	-0,524	-0,524	-0,522	-0,492	-0,330	0,020	0,562	1,069	2,849	4,114	10,064	72,696
B2_002	-0,705	-0,705	-0,705	-0,705	-0,705	-0,698	-0,690	-0,188	0,384	2,289	3,246	5,392	15,013
B2_013	-0,732	-0,732	-0,732	-0,716	-0,460	-0,100	-0,100	0,505	1,027	2,362	3,385	9,335	150,220
B2_003	-0,748	-0,748	-0,748	-0,736	-0,505	-0,265	-0,051	0,751	1,437	3,188	3,996	6,394	138,786
B2_000	-0,686	-0,686	-0,686	-0,669	-0,374	-0,227	0,081	0,856	1,473	3,124	3,986	8,141	112,485
B2_004	-0,982	-0,982	-0,982	-0,947	-0,522	-0,369	0,103	0,835	1,477	3,194	4,160	7,975	72,692
B2_005	-0,922	-0,922	-0,922	-0,892	-0,500	-0,339	0,092	0,785	1,426	2,980	3,781	7,049	82,859
B3_001	-0,184	-0,184	-0,184	-0,184	-0,184	-0,147	-0,109	0,023	0,108	0,399	0,732	3,290	141,072
B3_002	-0,182	-0,182	-0,182	-0,182	-0,182	-0,056	-0,051	0,005	0,073	0,342	0,541	1,566	90,774
B3_000	-0,159	-0,159	-0,159	-0,159	-0,159	-0,091	0,005	0,205	0,357	0,911	1,395	5,504	149,346
B3_003	-0,081	-0,081	-0,081	-0,081	-0,081	-0,047	-0,047	-0,047	-0,044	0,521	0,954	2,284	88,622
D1_001	-0,362	-0,362	-0,362	-0,362	-0,301	-0,214	-0,060	0,169	0,373	1,221	1,979	5,488	125,943
D1_002	-0,831	-0,831	-0,831	-0,831	-0,821	-0,538	-0,077	0,522	1,095	2,688	4,064	9,892	26,284
D1_003	-0,993	-0,993	-0,993	-0,993	-0,951	-0,685	0,034	0,732	1,209	2,696	3,643	6,652	59,592
D4_001	-0,077	-0,077	-0,077	-0,077	-0,073	-0,059	-0,020	0,080	0,150	0,384	0,745	3,707	133,901
D4_002	-0,141	-0,141	-0,137	-0,128	-0,115	-0,078	-0,001	0,109	0,190	0,725	1,447	7,753	161,342
D4_000	-0,118	-0,118	-0,114	-0,108	-0,097	-0,068	-0,011	0,086	0,182	0,641	1,184	6,445	141,722
D4_003	-0,092	-0,092	-0,092	-0,092	-0,092	-0,043	-0,011	0,041	0,088	0,253	0,442	1,432	188,412
D4_004	-0,180	-0,180	-0,174	-0,160	-0,088	-0,059	0,002	0,095	0,196	0,797	1,389	5,475	104,880
D4_006	-0,867	-0,867	-0,867	-0,867	-0,867	-0,413	-0,413	-0,413	-0,413	-0,413	-0,413	0,368	9,032

Fuente: Elaboración propia.

Dado que los indicadores analizados en ella son puntuaciones factoriales procedentes de la primera componente principal de los índices resumidos para cada variable en consideración, en estos cuadros, el valor mínimo se corresponde con el valor de aquellos hogares que no presentan incongruencia alguna para dicha variable, habiéndose considerado igualmente este valor para todos los hogares que no se tendrán en cuenta para la correspondiente corrección, por no haber pasado el correspondiente filtro de aplicación. Es por ello por lo que el valor mínimo puede copar los primeros percentiles de dichos cuadros.

Se observa en ellas también cómo hay un pequeño porcentaje de hogares (por encima del percentil del 99,9%) para los que los indicadores de congruencia se disparan, evidenciando la posibilidad de la existencia de casos muy atípicos que probablemente deberían haber sido aislados.

Análogamente, presentamos a continuación los respectivos cuadros para las partidas que intervienen en la Cuenta de Distribución Secundaria de la Rentas en los hogares: sus estadísticos básicos obtenidos para las distribuciones de valores de los indicadores presentados en el cuadro 4.15.a, una vez que fueron calculados cuando partíamos de los datos de las ECV base 2004 (cuadro 4.15.b) y de los de las ECV base 2013 (cuadro 4.15.c), observándose en ellos similares características.

Cuadro 4.15.a
INDICADORES ESPECÍFICOS
(CUENTA DE DISTRIBUCIÓN SECUNDARIA DE LA RENTA)

Partidas de la CRSH en la BDH		Variables BDH a Corregir / Calcular	Relaciones entre Variables usadas para calcular los Indicadores de Congruencia	Observaciones	Indicadores de Incongruencia	Filtro considerado para la aplicación de la corrección
Impuestos corrientes sobre la renta, el patrimonio, etc.	Impuestos sobre la renta	D_51_e_Bdh	D_51_e_Bdh (+) B_5n_Bdh		D5_001	(I=1)
	Otros impuestos corrientes	D_59_e_Bdh	D_59_e_Bdh (+) riquezaabr_eff	riquezabr_eff: Valor de RIQUEZA BRUTA (Total de Activos reales y Financieros)	D5_002	(I=1)
	Impuestos corrientes sobre la renta, el patrimonio, etc.	D_5_e_Bdh	D_5_e_Bdh = D_51_e_Bdh + D_59_e_Bdh			
Cotizaciones sociales	Cotizaciones sociales netas (empleos)	D_61_e_Bdh	D_61_e_Bdh (+) D_11_r_Bdh		D6_001	D_61_e_Bdh>0 or D_11_r_Bdh>0
	Cotizaciones sociales netas (Rec)	D_61_r_Bdh	D_61_r_Bdh (+) N TRABNEG_eff	N TRABNEG_eff: Número de trabajadores de los Negocios de la Persona de Referencia	D6_002	D_61_r_Bdh>0 or N TRABNEG_eff>0
	Cotizaciones sociales (empleos-recursos)	D_61_er_Bdh	D_61_er_Bdh = D_61_e_Bdh - D_61_r_Bdh			
Prestaciones sociales distintas de las transferencias sociales en especie	Prest. Soc. distintas de transf. soc. en espec (Rec)	D_62_r_Bdh	D_62_r_Bdh - CSIMEMP + INDNOVIDA1_B (+) PRESOC_B_m_eff	CSIMEMP: Cotizaciones Sociales imputadas a Cargo de los empleadores INDNOVIDA1_B: Indemnizaciones Brutas de Seguros de No Vida PRESOC_B_m_eff: Prestaciones sociales brutas (EFF)	D6_003	D_62_r_Bdh>0 or PRESOC_B_m_eff>0
	Prest. Soc. distintas de transf. soc. en espec (Emp)	D_62_e_Bdh	D_62_e_Bdh (+) N TRABNEG_eff	N TRABNEG_eff: Número de trabajadores de los Negocios de la Persona de Referencia	D6_004	D_62_e_Bdh>0 or N TRABNEG_eff>0
Otras transferencias corrientes	Indemnizaciones de seguro no vida (recursos)	D_72_r_Bdh	D_72_r_Bdh (+) PRIMNOVIDA_B	PRIMNOVIDA_B: Primas pagadas de Seguros de No Vida	D7_001	D_72_r_Bdh>0 or INDNOVIDA1_B>0
	Transferencias corrientes diversas (Recursos)	D_75_r_Bdh	D_75_r_Bdh (+) ICDOH_B_EFF	ICDOH_B_EFF: Transferencias corrientes brutas recibidas de otros hogares (EFF) OIRDOH_B: Rentas recibidas de Otros Hogares (ECV) P1_11d_eff=1: Otros miembros de fuera del hogar ayudan económicamente (EFF 2011)	D7_002	D_75_r_Bdh>0 or ICDOH_B_EFF>0 or OIRDOH_B>0 or (P1_11d_eff=1 and anoe=2011)
	Otras transferencias corrientes (Recursos)	D_7_r_Bdh	D_7_r_Bdh = D_72_r_Bdh + D_75_r_Bdh			
	Primas netas de seguro no vida	D_71_e_Bdh	D_71_e_Bdh (+) P5_23_eff	P5_23_eff: Pagos anuales en OTRAS FORMAS DE ASEGURAMIENTO (No Vida)	D7_003	D_71_e_Bdh>0 or P5_23_eff>0
	Transferencias corrientes diversas (empleos)	D_75_e_Bdh	D_75_e_Bdh (+) P9_4_EFF	P9_4_EFF: Dinero que se pasa a otras personas ajenas al hogar (EFF)	D7_004	D_75_e_Bdh>0 or P9_4_EFF>0
	Otras transferencias corrientes (Empleos)	D_7_e_Bdh	D_7_e_Bdh = D_71_e_Bdh + D_75_e_Bdh			
Renta disponible	Renta disponible neta	B_6n_Bdh	B_6n_Bdh = B_5n_Bdh - D_5_e_Bdh - D_61_er_Bdh + D_62_re_Bdh + D_7_r_Bdh - D_7_e_Bdh			
	Renta disponible bruta	B_6b_Bdh	B_6b_Bdh = B_5b_Bdh - D_5_e_Bdh - D_61_er_Bdh + D_62_re_Bdh + D_7_r_Bdh - D_7_e_Bdh			

Fuente: Elaboración propia a partir de la CNE, las ECV y las EFF.

Cuadro 4.15.b
INDICADORES ESPECÍFICOS ASOCIADOS A LAS PARTIDAS DE LA CUENTA DE
REDISTRIBUCIÓN DE RENTAS, CALCULADOS PARA LA BASE 2004

Indicadores	Mínimo	Percentiles											Máximo
		1	5	10	25	50	75	90	95	99	99,5	99,9	
D5_001	-0,226	-0,224	-0,212	-0,206	-0,179	-0,104	0,049	0,270	0,466	1,053	1,485	4,629	215,805
D5_002	-0,335	-0,325	-0,281	-0,226	-0,158	-0,057	0,055	0,161	0,244	0,885	1,747	7,785	96,319
D6_001	-0,205	-0,205	-0,205	-0,205	-0,198	-0,142	-0,100	0,097	0,336	1,217	1,548	2,604	182,552
D6_002	-0,142	-0,142	-0,142	-0,142	-0,142	-0,051	-0,051	0,066	0,227	0,655	0,875	2,818	85,621
D6_003	-0,388	-0,388	-0,383	-0,349	-0,276	-0,173	0,028	0,363	0,597	1,843	2,971	8,402	175,583
D6_004	-0,146	-0,146	-0,146	-0,146	-0,146	-0,051	-0,051	0,066	0,235	0,680	0,905	2,840	85,621
D7_001	-0,466	-0,466	-0,466	-0,466	-0,466	-0,408	-0,408	-0,408	-0,408	-0,408	-0,408	-0,053	5,066
D7_002	-0,892	-0,892	-0,892	-0,892	-0,892	-0,808	-0,808	-0,808	-0,808	0,828	1,278	3,148	10,741
D7_003	-0,033	-0,033	-0,033	-0,033	-0,033	-0,031	-0,031	-0,031	-0,031	0,664	1,299	3,313	90,393
D7_004	-0,708	-0,708	-0,708	-0,708	-0,708	-0,673	-0,673	-0,367	0,073	1,165	1,636	4,610	28,248
G_001	-0,544	-0,541	-0,521	-0,506	-0,445	-0,277	0,089	0,715	1,303	3,536	4,941	11,495	35,623

Fuente: Elaboración propia.

Cuadro 4.15.c
INDICADORES ESPECÍFICOS ASOCIADOS A LAS PARTIDAS DE LA CUENTA DE
REDISTRIBUCIÓN DE RENTAS, CALCULADOS PARA LA BASE 2013

Indicadores	Mínimo	Percentiles											Máximo
		1	5	10	25	50	75	90	95	99	99,5	99,9	
D5_001	-0,262	-0,259	-0,244	-0,218	-0,141	-0,081	0,034	0,212	0,374	0,972	1,608	4,833	170,697
D5_002	-0,384	-0,373	-0,323	-0,260	-0,192	-0,068	0,068	0,200	0,303	1,127	2,249	9,194	82,293
D6_001	-0,803	-0,803	-0,803	-0,803	-0,780	-0,524	-0,089	0,490	1,089	2,650	4,048	10,369	28,931
D6_002	-1,000	-1,000	-1,000	-1,000	-0,836	-0,514	0,014	0,506	1,049	3,155	5,262	9,179	13,977
D6_003	-0,347	-0,347	-0,339	-0,326	-0,273	-0,170	0,042	0,397	0,627	1,810	2,824	8,243	137,489
D6_004	-1,000	-1,000	-1,000	-1,000	-0,836	-0,503	0,013	0,494	1,021	3,181	5,289	9,394	13,976
D7_001	-0,545	-0,545	-0,545	-0,545	-0,545	-0,377	-0,377	-0,377	-0,377	-0,377	-0,377	0,078	4,758
D7_002	-0,842	-0,842	-0,842	-0,842	-0,842	-0,803	-0,803	-0,803	-0,803	0,746	1,095	2,633	10,932
D7_003	-0,032	-0,032	-0,032	-0,032	-0,032	-0,030	-0,030	-0,030	-0,030	0,497	1,236	3,928	98,654
D7_004	-0,655	-0,655	-0,655	-0,655	-0,655	-0,379	-0,379	-0,278	0,036	0,827	1,435	3,589	46,833

Fuente: Elaboración propia.

Asimismo, se observa en ellas cómo hay un pequeño porcentaje de hogares (por encima del percentil del 99,9%) para los que los indicadores de congruencia se disparan, evidenciando la posibilidad de la existencia de casos muy atípicos que probablemente deberían haber sido aislados.

4.1.2. Indicadores de incongruencia general

Finalmente, hemos calculado mediante la misma metodología algunos indicadores del nivel de incongruencia de las rentas declaradas para la renta disponible *per cápita* de los hogares, que nos servirán para evaluar el grado de ocultación global de rentas en cada hogar, independientemente de la casuística de su percepción.

En los siguientes cuadros se presentan tales indicadores (cuadro 4.16.a) y sus distribuciones: los estadísticos básicos obtenidos para las distribuciones de los valores de los indicadores presentados en el cuadro 4.16.a, una vez que fueron calculados cuando partíamos de los datos de las ECV base 2004 (cuadro 4.16.b) y de los de las ECV base 2013 (cuadro 4.16.c); los cuales nos permiten analizar someramente sus distribuciones.

Cuadro 4.16.a
INDICADORES GENERALES

Partidas de la CRSH en la BDH		Variables BDH a Corregir / Calcular	Relaciones entre Variables usadas para calcular los Indicadores de Congruencia	Observaciones	Indicadores de Incongruencia	Filtro considerado para la aplicación de la corrección
Renta per cápita	Renta disponible neta per cápita del hogar	B_6n_Bdh / nmh	B_6n_Bdh / nmh (+) GASTOH_epf / nmh	GASTOH_epf: Gasto Total del hogar (EPF) nmh: número de miembros del hogar	G_001	(I=1)
			B_6n_Bdh / nmh (+) GASTOH_eff / nmh	GASTOH_eff: Gasto Total del hogar (EFF) nmh: número de miembros del hogar	G_002	
			B_6n_Bdh / nmh (+) riquezazet / nmh	riquezazet_eff: riqueza neta del hogar (EFF) nmh: número de miembros del hogar	G_003	
			B_6n_Bdh / nmh (+) riquezazbr / nmh	riquezazbr_eff: riqueza bruta del hogar (EFF) nmh: número de miembros del hogar	G_004	
				Primera componente principal de G_001, G_002, G_003 y G_004	G_000	

Fuente: Elaboración propia a partir de la CNE, las EFF y las EPF.

Cuadro 4.16.b
INDICADORES GENERALES, CALCULADOS PARA LA BASE 2004

Indicadores	Mínimo	Percentiles											Máximo
		1	5	10	25	50	75	90	95	99	99,5	99,9	
G_001	-0,544	-0,541	-0,521	-0,506	-0,445	-0,277	0,089	0,715	1,303	3,536	4,941	11,495	35,623
G_002	-0,534	-0,530	-0,501	-0,486	-0,420	-0,235	0,135	0,678	1,163	2,847	3,940	7,921	150,290
G_003	-0,221	-0,219	-0,206	-0,187	-0,135	-0,080	0,028	0,193	0,340	0,898	1,538	5,668	199,758
G_004	-0,215	-0,214	-0,203	-0,186	-0,134	-0,081	0,024	0,190	0,342	0,909	1,525	5,746	196,184
G_000	-0,404	-0,390	-0,352	-0,329	-0,267	-0,141	0,074	0,389	0,690	1,914	2,864	7,526	172,646

Fuente: Elaboración propia a partir de las ECV base 2004.

Cuadro 4.16.c
INDICADORES GENERALES, CALCULADOS PARA LA BASE 2013

Indicadores	Mínimo	Percentiles											Máximo
		1	5	10	25	50	75	90	95	99	99,5	99,9	
G_001	-0,512	-0,509	-0,490	-0,475	-0,414	-0,251	0,098	0,649	1,182	3,098	4,599	10,234	69,142
G_002	-0,516	-0,513	-0,487	-0,445	-0,380	-0,212	0,117	0,604	1,054	2,596	3,806	8,748	147,713
G_003	-0,227	-0,225	-0,212	-0,194	-0,143	-0,086	0,026	0,194	0,345	0,907	1,591	5,822	111,260
G_004	-0,219	-0,218	-0,207	-0,189	-0,140	-0,085	0,022	0,188	0,342	0,902	1,561	5,665	111,968
G_000	-0,401	-0,386	-0,353	-0,330	-0,268	-0,143	0,071	0,380	0,676	1,871	2,839	7,878	88,421

Fuente: Elaboración propia a partir de las ECV base 2013.

Además, en estos indicadores se observan hogares por encima del percentil del 99,9% que disparan sus indicadores de incongruencia, registrando el resto de hogares diferentes niveles que evolucionan desde los mínimos, progresivamente, de forma bastante moderada.

5. Estimación combinada de la Distribución Personal de la Renta (DPR) y análisis de sus resultados

Una vez calculados los indicadores de incongruencia de las variables declaradas, podemos proceder a aplicar el método propuesto en este trabajo para calcular las ocultaciones imputables a cada caso y estimar consecuentemente la Distribución Personal de la Renta así corregida.

Puesto que nuestra intención fue aplicar el método a ambas bases para confrontar sus resultados y analizar la robustez del mismo, presentaremos por separados ambas estimaciones, ya que de los datos declarados en la base 2004 y la ocultación desvelada por la base 2013 para ellos, debemos estimar además los valores óptimos k que determinan objetivamente la propensión a la ocultación en cada partida, según hemos visto anteriormente, y que luego serán aplicados a los datos declarados en la base 2013.

5.1. Proceso de la Estimación combinada de la DPR (base 2004)

Antes de entrar a la aplicación del método, haremos un par de observaciones importantes sobre cómo actúan los indicadores de incongruencia en el mecanismo de corrección por ocultación de las rentas,

dependiendo de las cuatro situaciones que se pueden presentar, que resumimos en el cuadro 4.17.a, para el caso del indicador general de ocultación G_000.

Cuadro 4.17.a
**VARIACIÓN DE LA PENDIENTE DE LA TASA DE
OCULTACIÓN CUANDO AUMENTA G_000**

	Situación de Ocultación	Situación de Sobredeclaración
Variable tipo recurso	la pendiente se agudiza positivamente, sumándole una mayor cantidad ocultada al hogar a medida que crece el indicador de ocultación general G_000	la pendiente se agudiza negativamente y nos lleva al contrasentido de hacer decrecer más la variable (tipo recurso) y consecuentemente decrecer más sus rentas finales a medida que crece su indicador de ocultación general G_000
Variable tipo empleo	la pendiente se agudiza positivamente y nos lleva al contrasentido de hacer crecer más la variable (tipo empleo) y consecuentemente decrecer más sus rentas finales a medida que crece su indicador de ocultación general G_000	la pendiente se agudiza negativamente, restándole una mayor cantidad sobredeclarada al hogar en la variable (tipo empleo) y consecuentemente aumentando su renta final a medida que crece el indicador de ocultación general G_000

Fuente: Elaboración propia.

Con respecto de los indicadores de congruencia empleados, hay que señalar que como la tasa de ocultación del método propuesto toma siempre el mismo signo que la ocultación agregada a corregir, entonces la corrección que se realizará será tanto más intensa (positiva o negativamente) cuanto mayor sea el indicador de incongruencia específico directamente relacionado con el nivel de ocultación de la variable de ingresos de referencia en cada caso.

Sin embargo, el crecimiento del indicador de incongruencia general, empleado para estimar la tasa de ocultación, producirá pendientes más acusadas (positivas o negativas) de dicha tasa, independientemente de que la variable a corregir se refiera a recursos o empleos de la CRSH, lo que debemos solventar.

Para corregir este efecto indeseable, basta con cambiar el signo del indicador de incongruencia para que actúe en la dirección conveniente. En resumen, el indicador que finalmente emplearemos para estimar la tasa de ocultación, dependiendo de la situación, será elegido de acuerdo con el siguiente cuadro 4.17.b; y, consecuentemente, los indicadores y filtros empleados para estimar los valores de k y corregir cada tipo de renta para los datos de la base 2004 se muestran en los cuadros 4.18.a y cuadro 4.18.b.

Cuadro 4.17.b
ELECCIÓN DEL INDICADOR PARA ESTIMAR LA PENDIENTE
DE LA TASA DE OCULTACIÓN

	Situación de Ocultación	Situación de Sobredeclaración
Variable tipo recurso	G_000	- G_000
Variable tipo empleo	- G_000	G_000

Fuente: Elaboración propia.

Cuadro 4.18.a
ESPECIFICACIONES DE LOS MODELOS DE CORRECCIÓN DE LAS PARTIDAS DE LA
CUENTA DE ASIGNACIÓN DE LA RENTA PRIMARIA, PARA LA BASE 2004

Partidas de la CRSH en la BDH		Variables BDH a Corregir / Calcular	Ocultación a Corregir	Filtro considerado para la aplicación de la corrección	Indicador de Incongruencia directamente relacionado con el nivel de la ocultación	Indicador de incongruencia empleado para estimar la tasa de ocultación
Excedente de Explotación	Producción relacionada con el Alquiler imputado Vivienda Principal, si el Hogar la ocupa en propiedad	AIVPP_P	2008: Ocultación O_AIVPP_P2 2011: Ocultación	regteneviv=1	B2_001	2008: G_000 2011: G_000
	Intereses pagados de préstamos hipotecarios	AIVPP_CI_JPPH	2008: Sobredeclaración O_AIVPP_CI_JPPH 2011: Sobredeclaración	regteneviv=2	B2_002	2008: G_000 2011: G_000
	Servicios de Seguros y Planes de Pensiones	AIVPP_CI_Seg	2008: Sobredeclaración O_AIVPP_CI_Seg 2011: Ocultación	regteneviv=2	B2_000	2008: G_000 2011: - G_000
	Otros Impuestos netos sobre la Producción relacionados con el AIVPP	AIVPP_OIN	2008: Sobredeclaración O_AIVPP_OIN 2011: Ocultación	regteneviv=1	B2_004	2008: G_000 2011: - G_000
	Consumo de Capital Fijo del Excedente de Explotación	CCFEE	2008: Sobredeclaración O_CCFEE 2011: Ocultación	regteneviv=1	B2_005	2008: G_000 2011: - G_000
Renta mixta	Renta Mixta Neta	B_3n_Bdh	2008: Ocultación O_B_3n 2011: Ocultación	B_3n_Bdh<0 or (RMIX_B_EFF + IMALQPRO_EFF<0)	B3_000	2008: G_000 2011: G_000
	Consumo de Capital Fijo de la Renta Mixta	CCFRM	2008: Ocultación O_CCFRM 2011: Ocultación	B_3n_Bdh<0 or (RMIX_B_EFF + IMALQPRO_EFF<0)	B3_003	2008: - G_000 2011: - G_000
Remuneración de Asalariados	Sueldos y Salarios	D_11_r_Bdh	2008: Ocultación O_D_11_r 2011: Ocultación	D_11_r_Bdh>0 or (RMCA_B_EFF+RNMCA_B_EFF>0)	D1_001	2008: G_000 2011: G_000
	Cotizaciones sociales efectivas a cargo de empleadores	D_121_r_Bdh	2008: Ocultación O_D_121_r 2011: Ocultación	D_11_r_Bdh>0 or (RMCA_B_EFF+RNMCA_B_EFF>0)	D1_002	2008: G_000 2011: G_000
	Cotizaciones sociales imputadas a cargo de empleadores	D_122_r_Bdh	2008: Ocultación O_D_122_r 2011: Ocultación	D_11_r_Bdh>0 or (RMCA_B_EFF+RNMCA_B_EFF>0)	D1_003	2008: G_000 2011: G_000
Rentas de la propiedad	Intereses (recursos) + Rentas distribuidas de las sociedades (recursos)	D_4142_r_Bdh	2008: Ocultación O_D_4142_r 2011: Ocultación	D_4142_r_Bdh>0 or IMCAP_B_EFF>0 or actfinanc_eff>0	D4_000	2008: G_000 2011: G_000
	Intereses (empleos)	D_41_e_Bdh	2008: Ocultación O_D_41_e 2011: Ocultación	D_41_e_Bdh>0 or devotras_eff>0	D4_003	2008: - G_000 2011: - G_000
	Otras rentas de inversión	D_44_r_Bdh	2008: Ocultación O_D_44_r 2011: Ocultación	D_44_r_Bdh>0 or actfinanc_eff>0	D4_004	2008: G_000 2011: G_000
	Rentas de la tierra (rec-emp)	D_45_re	2008: Ocultación O_D_45_re 2011: Ocultación	D_45_re>0 or IMALQTER_EFF>0	D4_006	2008: G_000 2011: G_000

Fuente: Elaboración propia a partir de la CNE, las ECV y las EFF.

Cuadro 4.18.b
ESPECIFICACIONES DE LOS MODELOS DE CORRECCIÓN DE LAS PARTIDAS DE LA CUENTA DE DISTRIBUCIÓN SECUNDARIA DE LA RENTA, PARA LA BASE 2004

Partidas de la CRSH en la BDH		Variables BDH a Corregir / Calcular	Ocultación a Corregir	Filtro considerado para la aplicación de la corrección	Indicador de Incongruencia directamente relacionado con el nivel de la ocultación	Indicador de incongruencia empleado para estimar la tasa de ocultación
Impuestos corrientes sobre la renta, el patrimonio, etc.	Impuestos sobre la renta	D_51_e_Bdh	2008: Ocultación O_D_51_e 2011: Ocultación	(1=1)	D5_001	2008: - G_000 2011: - G_000
	Otros impuestos corrientes	D_59_e_Bdh	2008: Ocultación O_D_59_e 2011: Ocultación	(1=1)	D5_002	2008: - G_000 2011: - G_000
Cotizaciones sociales	Cotizaciones sociales netas (empleos)	D_61_e_Bdh	2008: Ocultación O_D_61_e 2011: Ocultación	D_61_e_Bdh>0 or D_1_r_Bdh>0	D6_001	2008: -G_000 2011: - G_000
	Cotizaciones sociales netas (Rec)	D_61_r_Bdh	2008: Ocultación O_D_61_r 2011: Ocultación	D_61_r_Bdh>0 or N_TRABNEG_eff>0	D6_002	2008: G_000 2011: G_000
Prestaciones sociales distintas de las transferencias sociales en especie	Prest. Soc. distintas de transf. soc. en espec (Rec)	D_62_r_Bdh	2008: Ocultación O_D_62_r 2011: Ocultación	D_62_r_Bdh>0 or PRESOC_B_m_eff>0	D6_003	2008: G_000 2011: G_000
	Prest. Soc. distintas de transf. soc. en espec (Emp)	D_62_e_Bdh	2008: Ocultación O_D_62_e 2011: Ocultación	D_62_e_Bdh>0 or N_TRABNEG_eff>0	D6_004	2008: - G_000 2011: - G_000
Otras transferencias corrientes	Indemnizaciones de seguro no vida (recursos)	D_72_r_Bdh	2008: Ocultación O_D_72_r 2011: Ocultación	D_72_r_Bdh>0 or INDNOVIDA1_B>0	D7_001	2008: G_000 2011: G_000
	Transferencias corrientes diversas (Recursos)	D_75_r_Bdh	2008: Ocultación O_D_75_r 2011: Ocultación	D_75_r_Bdh>0 or TCDOH_B_EFF>0 or OTRDOH_B>0 or (P1_11d_eff=1 and aage=2011)	D7_002	2008: G_000 2011: G_000
	Primas netas de seguro no vida	D_71_e_Bdh	2008: Ocultación O_D_71_e 2011: Ocultación	D_71_e_Bdh>0 or P5_23_eff>0	D7_003	2008: - G_000 2011: - G_000
	Transferencias corrientes diversas (empleos)	D_75_e_Bdh	2008: Ocultación O_D_75_e 2011: Ocultación	D_75_e_Bdh>0 or P9_4_EFF>0	D7_004	2008: - G_000 2011: - G_000

Fuente: Elaboración propia a partir de la CNE, las ECV y las EFF.

5.1.1. La progresividad estimada “k” de las tasas de ocultación

Como hemos dicho anteriormente, para estimar la progresividad de las tasas de ocultación (k) es preciso disponer de información previa acerca de cómo suele ser el comportamiento de los hogares ante el fenómeno de la ocultación de rentas, según la tipología de éstas. Así pues, y mediante la metodología de fusión estadística descrita anteriormente, hemos procedido a obtener la ocultación desvelada por los registros administrativos empleados para la ECV base 2013, con respecto de las declaraciones realizadas en las ECV b2004 por los hogares, para cada hogar que efectivamente ocultó.

Con estas ocultaciones de cada hogar en cada partida de renta, y considerando los indicadores de incongruencia y especificaciones recogidos en los cuadros 4.18.a y 4.18.b, mediante la metodología expuesta en el apartado 2.4 de este capítulo, se procedió a estimar los parámetros “k”, inicialmente indeterminados para la expresión lineal de la tasa de ocultación implícita en el modelo, para los años de referencia de los ingresos considerados en este estudio (2008 y 2011), presentándose las estimaciones resultantes en el cuadro 4.19.

Observamos en esos resultados que en una mayoría de los casos (30 situaciones), la constante “k” toma el valor cero, por lo que en estas situaciones, la solución que más se adapta a las correcciones realizadas entre la base 2004 y la base 2013 implica que la tasa de corrección sobre los niveles de incongruencia detectados sea constante (corrección directamente proporcional al indicador específico de incongruencia, sin ninguna otra matización).

En menor medida, la tasa de ocultación se manifiesta realmente progresiva en otras 18 situaciones, de las que en 11 de ellas la progresividad es la máxima permitida por el método ($k=1$).

Cuadro 4.19
VALORES ESTIMADOS PARA LA PROGRESIVIDAD “K”

	2008	2011
AIVPP_P	0,000906	0,000000
AIVPP_CI_IPPH	0,000000	0,000000
AIVPP_CI_Seg	0,067459	0,000000
AIVPP_OIN	0,000000	1,000000
CCFEE	0,000000	1,000000
B_3n	0,000000	0,000000
CCFRM	0,000000	0,000000
D_11_r	0,000000	0,000000
D_121_r	0,000000	0,177737
D_122_r	0,000000	0,000000
D_4142_r	0,000000	0,016009
D_41_e	1,000000	0,000000
D_44_r	0,000000	0,000000
D_45_re	0,000000	0,000000
D_51_e	1,000000	0,000000
D_59_e	1,000000	1,000000
D_61_e	1,000000	0,000000
D_61_r	0,000000	0,036913
D_62_r	0,000000	0,000000
D_62_e	1,000000	0,000000
D_72_r	0,304770	1,000000
D_75_r	0,000000	0,162541
D_71_e	0,000000	1,000000
D_75_e	1,000000	0,000000

Fuente: Elaboración propia.

Siendo éstos valores de “k” los óptimos de mínimos cuadrados obtenidos a partir de las ocultaciones desveladas por la comparación de las dos bases de las ECV, y pese a la lógica de decisión de modelar la tasa de ocultación con respecto de los índices de incongruencia de forma lineal, el hecho de que en sólo siete de las situaciones el valor de k diera valores intermedios entre 0 y 1, nos hace pensar en la conveniencia de cuestionarnos si, efectivamente, la forma lineal propuesta para la tasa de ocultación es la más adecuada o no. Pero la investigación de esta potencial mejora trasciende el horizonte de esta investigación, quedando pospuesta para una posible futura línea de investigación.

Y, una vez estimados los valores óptimos de k, se procedió a aplicar el método de corrección a todas las partidas de los cuadros 4.18.a y 4.18.b, secuencialmente en tres etapas: primero las partidas señaladas en negro; y, una vez corregidas estas, las marcadas en color naranja en otras dos etapas, para las partidas que dependían de partidas corregidas en una etapa anterior.

El programa informático que realiza tanto la estimación del parámetro k como la obtención previa de las ocultaciones desveladas por la base 2013 de las ECV para cada caso, se presenta en el Anexo 4. Y el programa que realiza la corrección de todas las partidas, previo cálculo concreto de los pertinentes indicadores de incongruencia, se presenta en el Anexo 5.

5.2. Proceso de la Estimación combinada de la DPR (base 2013)

Con objeto de valorar la robustez de las estimaciones que el método provee, hemos procedido análogamente con los datos de la ECV base 2013, pero en este caso (en el que, a diferencia del anterior, no disponemos de información de contraste para analizar la ocultación de los hogares) hemos empleado las mismas propensiones a la ocultación en cada partida (parámetros k) que hemos estimado por mínimos cuadrados (como ya se ha dicho para los datos base 2004).

Los indicadores y filtros empleados para corregir cada tipo de renta para los datos de la base 2013 se muestran en los siguientes cuadros 4.20.a y 4.20.b.

Cuadro 4.20.a
ESPECIFICACIONES DE LOS MODELOS DE CORRECCIÓN DE LAS PARTIDAS DE LA CUENTA DE ASIGNACIÓN DE LA RENTA PRIMARIA, PARA LA BASE 2013

Partidas de la CRSH en la BDH		Variables BDH a Corregir / Calcular	Ocultación a Corregir	Filtro considerado para la aplicación de la corrección	Indicador de Incongruencia directamente relacionado con el nivel de la ocultación	Indicador de incongruencia empleado para estimar la tasa de ocultación
Excedente de Explotación	Producción relacionada con el Alquiler imputado Vivienda Principal, si el Hogar la ocupa en propiedad	AIVPP_P	2008: Ocultación O_AIVPP_P2 2011: Ocultación	regteneviv=1	B2_001	2008: G_000 2011: G_000
	Intereses pagados de préstamos hipotecarios	AIVPP_CI_IPPH	2008: Sobredeclaración O_AIVPP_CI_IPPH 2011: Sobredeclaración	regteneviv=2	B2_002	2008: G_000 2011: G_000
	Servicios de Seguros y Planes de Pensiones	AIVPP_CI_Seg	2008: Sobredeclaración O_AIVPP_CI_Seg 2011: Ocultación	regteneviv=2	B2_000	2008: G_000 2011: - G_000
	Otros Impuestos netos sobre la Producción relacionados con el AIVPP	AIVPP_OIN	2008: Sobredeclaración O_AIVPP_OIN 2011: Ocultación	regteneviv=1	B2_004	2008: G_000 2011: - G_000
	Consumo de Capital Fijo del Excedente de Explotación	CCFEE	2008: Sobredeclaración O_CCFEE 2011: Ocultación	regteneviv=1	B2_005	2008: G_000 2011: - G_000
Renta mixta	Renta Mixta Neta	B_3n_Bdh	2008: Ocultación O_B_3n 2011: Ocultación	B_3n_Bdh<0 or (RMIX_B_EFF + IMALQPRO_EFF<0)	B3_000	2008: G_000 2011: G_000
	Consumo de Capital Fijo de la Renta Mixta	CCFRM	2008: Ocultación O_CCFRM 2011: Ocultación	B_3n_Bdh<0 or (RMIX_B_EFF + IMALQPRO_EFF<0)	B3_003	2008: - G_000 2011: - G_000
Remuneración de Asalariados	Sueldos y Salarios	D_11_r_Bdh	2008: Ocultación O_D_11_r 2011: Ocultación	D_11_r_Bdh>0 or (RMCA_B_EFF+RMCA_B_EFF>0)	D1_001	2008: G_000 2011: G_000
	Cotizaciones sociales efectivas a cargo de empleadores	D_121_r_Bdh	2008: Sobredeclaración O_D_121_r 2011: Sobredeclaración	D_11_r_Bdh>0 or (RMCA_B_EFF+RMCA_B_EFF>0)	D1_002	2008: - G_000 2011: - G_000
	Cotizaciones sociales imputadas a cargo de empleadores	D_122_r_Bdh	2008: Sobredeclaración O_D_122_r 2011: Sobredeclaración	D_11_r_Bdh>0 or (RMCA_B_EFF+RMCA_B_EFF>0)	D1_003	2008: - G_000 2011: - G_000
Rentas de la propiedad	Intereses (recursos) + Rentas distribuidas de las sociedades (recursos)	D_4142_r_Bdh	2008: Ocultación O_D_4142_r 2011: Ocultación	D_4142_r_Bdh>0 or IMCAP_B_EFF>0 or actfinanc_eff>0	D4_000	2008: G_000 2011: G_000
	Intereses (empleos)	D_41_e_Bdh	2008: Ocultación O_D_41_e 2011: Ocultación	D_41_e_Bdh>0 or deutos_eff>0	D4_003	2008: - G_000 2011: - G_000
	Otras rentas de inversión	D_44_r_Bdh	2008: Ocultación O_D_44_r 2011: Ocultación	D_44_r_Bdh>0 or actfinanc_eff>0	D4_004	2008: G_000 2011: G_000
	Rentas de la tierra (rec-emp)	D_45_re	2008: Sobredeclaración O_D_45_re 2011: Ocultación	D_45_re>0 or IMALQTER_EFF>0	D4_006	2008: - G_000 2011: G_000

Fuente: Elaboración propia a partir de la CNE, las ECV y las EFF.

Cuadro 4.20.b
ESPECIFICACIONES DE LOS MODELOS DE CORRECCIÓN DE LAS PARTIDAS DE LA CUENTA DE DISTRIBUCIÓN SECUNDARIA DE LA RENTA, PARA LA BASE 2013

Partidas de la CRSH en la BDH		Variables BDH a Corregir / Calcular	Ocultación a Corregir	Filtro considerado para la aplicación de la corrección	Indicador de Incongruencia directamente relacionado con el nivel de la ocultación	Indicador de incongruencia empleado para estimar la tasa de ocultación
Impuestos corrientes sobre la renta, el patrimonio, etc.	Impuestos sobre la renta	D_51_e_Bdh	2008: Ocultación O_D_51_e 2011: Ocultación	(1=1)	D5_001	2008: -G_000 2011: -G_000
	Otros impuestos corrientes	D_59_e_Bdh	2008: Ocultación O_D_59_e 2011: Ocultación	(1=1)	D5_002	2008: -G_000 2011: -G_000
Cotizaciones sociales	Cotizaciones sociales netas (empleos)	D_61_e_Bdh	2008: Sobredeclaración O_D_61_e 2011: Sobredeclaración	D_61_e_Bdh>0 or D_1_r_Bdh>0	D6_001	2008: G_000 2011: G_000
	Cotizaciones sociales netas (Rec)	D_61_r_Bdh	2008: Sobredeclaración O_D_61_r 2011: Sobredeclaración	D_61_r_Bdh>0 or N TRABNEG_eff>0	D6_002	2008: -G_000 2011: -G_000
Prestaciones sociales distintas de las transferencias sociales en especie	Prest. Soc. distintas de transf. soc. en espec (Rec)	D_62_r_Bdh	2008: Ocultación O_D_62_r 2011: Ocultación	D_62_r_Bdh>0 or PRESOC_B_m_eff>0	D6_003	2008: G_000 2011: G_000
	Prest. Soc. distintas de transf. soc. en espec (Emp)	D_62_e_Bdh	2008: Sobredeclaración O_D_62_e 2011: Sobredeclaración	D_62_e_Bdh>0 or N TRABNEG_eff>0	D6_004	2008: G_000 2011: G_000
Otras transferencias corrientes	Indemnizaciones de seguro no vida (recursos)	D_72_r_Bdh	2008: Ocultación O_D_72_r 2011: Ocultación	D_72_r_Bdh>0 or INDNOVIDA1_B>0	D7_001	2008: G_000 2011: G_000
	Transferencias corrientes diversas (Recursos)	D_75_r_Bdh	2008: Ocultación O_D_75_r 2011: Ocultación	D_75_r_Bdh>0 or TCDOH_B_EFF>0 or OTRDOH_B>0 or (P1_11d_eff=1 and anoe=2011)	D7_002	2008: G_000 2011: G_000
	Primas netas de seguro no vida	D_71_e_Bdh	2008: Ocultación O_D_71_e 2011: Ocultación	D_71_e_Bdh>0 or P5_23_eff>0	D7_003	2008: -G_000 2011: -G_000
	Transferencias corrientes diversas (empleos)	D_75_e_Bdh	2008: Ocultación O_D_75_e 2011: Ocultación	D_75_e_Bdh>0 or P9_4_EFF>0	D7_004	2008: -G_000 2011: -G_000

Fuente: Elaboración propia a partir de la CNE, las ECV y las EFF.

Como puede observarse, estos cuadros son muy parecidos a los cuadros 4.18.a y 4.18.b, respectivamente, pero no son idénticos. Por un lado, hay diferencias que se encuentran en la columna "Ocultación a Corregir", ya que la corrección a realizar deberá ser por ocultación o por sobredeclaración dependiendo de los agregados de cada partida en cada año y su posición relativa con respecto de los de CNE. Y, por otro lado, consecuentemente con lo anterior y con lo resumido en los cuadros 4.17.a y 4.17.b, el indicador de incongruencia empleado para estimar la progresividad de la tasa de ocultación (k) será +G_000 o -G_000, dependiendo de la situación.

5.3. Las estimaciones obtenidas

5.3.1. Distribuciones personales la renta para el año 2008

El cuadro 4.21.a presenta, para el año 2008, la comparación de los agregados de cada partida de la CRSH proporcionado por la CNE (base 2010) y los contruidos en cada una de las cuatro situaciones que hemos considerado en este estudio:

- Contruidos directamente a partir de los datos de las ECV base 2004 (columna BDH (b2004)).

- Construidos directamente a partir de los datos de las ECV base 2013 (columna BDH (b2013)).
- Construidos a partir de los datos de las ECV base 2004, las EPF y las EFF, corregidos mediante la aplicación de la metodología propuesta en este trabajo (columna BDH (b2004) Corregida).
- Construidos a partir de los datos de las ECV base 2013, las EPF y las EFF, corregidos mediante la aplicación de la metodología propuesta en este trabajo (columna BDH (b2013) Corregida).

Cuadro 4.21.a
**CONGRUENCIA DE LA DISTRIBUCIÓN PERSONAL DE LA RENTA
CORREGIDA CON CNE (2008)**

		BDH (b2004)	BDH (b2013)	CNE	BDH (b2004) Corregida	BDH (b2013) Corregida
Npers	Número de personas	45.870.296	45.870.296	45.983.168	45.983.169	45.983.169
B_2n	Excedente Neto de Explotación	60.192.213.232	59.990.047.196	31.734.000.000	31.734.000.000	31.734.000.000
B_3n	Renta mixta neta	37.965.307.827	67.638.559.753	96.589.000.000	96.589.000.000	96.589.000.000
B_2n3n	Excedente de explotación neto / Renta mixta neta	98.157.521.059	127.628.606.949	128.323.000.000	128.323.000.000	128.323.000.000
D_11_r	Sueldos y Salarios	350.613.064.932	388.944.188.599	442.319.000.000	442.319.000.000	442.319.000.000
D_121_r	Cotizaciones sociales efectivas a cargo de empleadores	84.146.232.214	99.798.470.083	99.045.798.352	99.045.798.352	99.045.798.352
D_122_r	Cotizaciones sociales imputadas a cargo de empleadores			19.152.201.648	19.152.201.648	19.152.201.648
D_12_r	Cotizaciones sociales a cargo del empleador	84.146.232.214	99.798.470.083	118.198.000.000	118.198.000.000	118.198.000.000
D_1_r	Remuneración de Asalariados	434.759.297.146	488.742.658.682	560.517.000.000	560.517.000.000	560.517.000.000
D_41_e	Intereses (empleos)			40.504.000.000	40.504.000.000	40.504.000.000
D_4142_re	Saldo (Rec-Emp) de Intereses, Dividendos y Ganacias de Inversiones de Capital			11.701.000.000	11.701.000.000	11.701.000.000
D_44_r	Rentas de la propiedad atribuidas a los asegurados (recursos)			25.016.000.000	25.016.000.000	25.016.000.000
D_45_re	Rentas de la tierra (rec-emp)	7.535.925.685	8.819.261.808	79.000.000	79.000.000	79.000.000
D_4_re	Rentas de la propiedad (rec-emp)	12.430.664.601	31.180.885.167	36.796.000.000	36.796.000.000	36.796.000.000
B_5n	Saldo de rentas primarias netas	545.347.482.806	647.552.150.798	725.636.000.000	725.636.000.000	725.636.000.000
D_51_e	Impuestos sobre la renta	68.757.130.504	90.684.483.713	79.865.000.000	79.865.000.000	79.865.000.000
D_59_e	Otros impuestos corrientes	691.365.008	712.068.463	4.907.000.000	4.907.000.000	4.907.000.000
D_5_e	Impuestos corrientes sobre la renta, el patrimonio, etc.	69.448.495.513	91.396.552.176	84.772.000.000	84.772.000.000	84.772.000.000
D_61_r	Cotizaciones sociales (recursos)			308.000.000	308.000.000	308.000.000
D_61_e	Cotizaciones sociales (empleos)	84.146.232.214	99.798.470.083	163.579.000.000	163.579.000.000	163.579.000.000
D_62_r	Prest. Soc. distintas de transf. soc. en espec (Rec)	113.851.048.059	126.662.474.249	154.572.000.000	154.572.000.000	154.572.000.000
D_62_e	Prest. Soc. distintas de transf. soc. en espec (Emp)			369.000.000	369.000.000	369.000.000
D_71_e	Primas netas de seguro no vida			13.461.000.000	13.461.000.000	13.461.000.000
D_72_r	Indemnizaciones de seguro no vida (recursos)			13.580.000.000	13.580.000.000	13.580.000.000
D_75_r	Transferencias corrientes diversas (Recursos)	2.273.759.577	2.263.734.161	37.985.000.000	37.985.000.000	37.985.000.000
D_75_e	Transferencias corrientes diversas (empleos)	4.200.439.344	4.355.667.692	41.133.000.000	41.133.000.000	41.133.000.000
D_7_r	Otras transferencias corrientes (Recursos)	2.273.759.577	2.263.734.161	51.565.000.000	51.565.000.000	51.565.000.000
D_7_e	Otras transferencias corrientes (Empleos)	4.200.439.344	4.355.667.692	54.594.000.000	54.594.000.000	54.594.000.000
B_6n	Renta disponible Neta	503.677.123.370	580.927.669.257	628.767.000.000	628.767.000.000	628.767.000.000
K_1	Consumo de Capital Fijo			45.936.000.000	45.936.000.000	45.936.000.000

Fuente: Elaboración propia a partir de los datos de las ECV y la CNE base 2010.

Como podemos observar en él, los agregados para las distribuciones corregidas en este año, resultantes de la aplicación del procedimiento propuesto tanto a los datos de la base 2004 como a los datos de la base 2013, respetan escrupulosamente los agregados marcados por CNE que hemos considerado como referencia.

La forma en que las estimaciones resultantes logran esta congruencia agregada con CNE, así como el grado de robustez de las estimaciones que el método propuesto proporciona, podemos estudiarla someramente en los cuadros 4.21.b y 4.21.c.

En el cuadro 4.21.b se recogen comparativamente los principales estadísticos de la distribución personal de la renta disponible (per cápita), en cada una de las cuatro situaciones que hemos considerados en este estudio:

- Distribución obtenida directamente a partir de la variable "vhRentaAla" de las ECV base 2004 (Renta disponible total del hogar en el año anterior al de

la entrevista, incluyendo el alquiler imputado), dividiéndola por el número de miembros del hogar (columna ECV b2004 (vhRentaAla pc)).

- Distribución obtenida directamente a partir de la variable "vhRentaAla" de las ECV base 2013 (Renta disponible total del hogar en el año anterior al de la entrevista, incluyendo el alquiler imputado), dividiéndola por el número de miembros del hogar (columna ECV b2013 (vhRentaAla pc)).
- Construidos a partir de los datos de las ECV base 2004, las EPF y las EFF, corregidos mediante la aplicación de la metodología propuesta en este trabajo (columna Corregida a partir de la b2004).
- Construidos a partir de los datos de las ECV base 2013, las EPF y las EFF, corregidos mediante la aplicación de la metodología propuesta en este trabajo (columna Corregida a partir de la b2013)).

Cuadro 4.21.b
DISTRIBUCIÓN DE LA RENTA DISPONIBLE NETA PER CÁPITA (2008)

	ECV b2004 (vhRentaAla pc)	ECV b2013 (vhRentaAla pc)	Corregida a partir de la b2004	Corregida a partir de la b2013
Nº de casos considerados	13.360	13.360	61.645	62.804
Hogares representados	17.279.282	17.279.282	17.324.391	17.324.391
Nº de personas representadas	45.870.296	45.870.296	45.983.169	45.983.169
Tamaño medio de los hogares	2,65	2,65	2,65	2,65
Mínimo	-75.001	-56.535	-532.872	-1.098.126
Percentil 5	2.730	3.394	2.882	3.236
Percentil 10	4.160	4.718	4.932	4.889
Percentil 25	6.635	7.392	7.744	7.527
Percentil 50	9.792	11.124	11.411	11.320
Percentil 75	14.025	15.995	16.555	16.864
Percentil 90	19.339	22.288	24.103	24.462
Percentil 95	24.068	28.110	30.567	31.210
Máximo	133.853	400.073	2.040.961	1.730.273
Media	11.103	12.790	13.674	13.674
Desviación Típica	7.219	8.910	16.641	14.169
Coeficiente de Variación de Pearson	65,0%	69,7%	121,7%	103,6%
Total agregado (Millones €)	509.313	586.681	628.767	628.767
Recursos acumulados hasta el Percentil 5	0,28%	0,66%	-0,52%	-0,69%
Recursos acumulados hasta el Percentil 10	1,83%	2,26%	0,97%	0,85%
Recursos acumulados hasta el Percentil 25	9,23%	9,47%	8,06%	7,75%
Recursos acumulados hasta el Percentil 50	27,65%	27,53%	25,54%	24,89%
Recursos acumulados hasta el Percentil 75	54,13%	53,57%	50,60%	50,10%
Recursos acumulados hasta el Percentil 90	76,09%	75,28%	72,21%	72,03%
Recursos acumulados hasta el Percentil 95	85,68%	84,97%	82,04%	82,03%
Índice de Gini	0,3274	0,3305	0,3722	0,3797

Fuente: Elaboración propia a partir de los datos de las ECV y la CNE base 2010.

Comparando las distribuciones presentadas en el cuadro 4.21.b para la renta disponible per cápita en los hogares en 2008, puede verse que las correcciones más acusadas se han realizado sobre ambas colas de las distribuciones originales, entre el mínimo y el percentil 5, y entre el percentil 95 y el máximo; lo que produce un cierto incremento en la dispersión de las distribuciones corregidas con respecto de las originales, pasando el Coeficiente de variación de Pearson de estar en torno al 70% en las distribuciones originales según las ECV, a superar el 100% de la media.

También la distribución de recursos se ve desplazada especialmente hacia las rentas mayores que la mediana, agudizándose su efecto a partir del

percentil 95 de la distribución que pasa de acaparar aproximadamente un 15% de los recursos según las ECV, a aproximadamente un 18% después de la corrección.

Pero si comparamos las dos últimas columnas de este cuadro, construidas para comprobar la robustez de las estimaciones que el método proporciona, ya que parten de microdatos sustancialmente diferentes (base 2004 y base 2013), observamos que las distribuciones obtenidas por ambas aplicaciones del método de corrección son muy parecidas, quizás con la excepción de sus valores más extremos que parecen haber sido muy especialmente corregidos por el procedimiento y que han producido coeficientes de variación de Pearson algo diferentes (121,7% frente al 103,6%). Sin embargo, coinciden sus medias, los percentiles apenas difieren (en el peor de los casos, unos pocos cientos de euros en los extremos), y los recursos acumulados por tramos se muestran igualmente bastante estables (con variaciones inferiores en todo caso a medio punto de porcentaje⁷⁰).

Análogamente, en el cuadro 4.21.c se recogen comparativamente los principales estadísticos de la distribución personal de la renta personal equivalente disponible (per cápita), para 2008, de acuerdo con la escala de equivalencia de la OCD modificada, en cada una de las cuatro situaciones que hemos considerados en este estudio:

- Distribución obtenida directamente a partir de la variable "vhRentaAla" de las ECV base 2004 (Renta disponible total del hogar en el año anterior al de la entrevista, incluyendo el alquiler imputado), dividiéndola por el número de unidades de consumo, según la escala UCDE modificada (columna ECV b2004 (vhRentaAla pq2)).
- Distribución obtenida directamente a partir de la variable "vhRentaAla" de las ECV base 2013 (Renta disponible total del hogar en el año anterior al de la entrevista, incluyendo el alquiler imputado), dividiéndola por el número de unidades de consumo, según la escala UCDE modificada (columna ECV b2013 (vhRentaAla pq2)).
- Construidos a partir de los datos de las ECV base 2004, las EPF y las EFF, corregidos mediante la aplicación de la metodología propuesta en este trabajo (columna Corregida a partir de la b2004).
- Construidos a partir de los datos de las ECV base 2013, las EPF y las EFF, corregidos mediante la aplicación de la metodología propuesta en este trabajo (columna Corregida a partir de la b2013)).

Se aprecian comportamientos muy parecidos a los ya observados en el cuadro 4.21.b para la renta disponible per cápita en los hogares en 2008, Puede verse en este caso, que las correcciones más acusadas se materializan a partir de la mediana de la distribución, haciendo perder peso relativo a las clases bajas en el reparto total de recursos. Se produce, de forma análoga que en el caso anterior, un cierto incremento, en la dispersión de las distribuciones corregidas con respecto de las originales, pasando el Coeficiente de variación de Pearson de estar en torno al 60% en las

⁷⁰ Como ya comentamos anteriormente, será preciso estudiar en el futuro el efecto que sobre los datos muy atípicos está teniendo el método tal como ha sido propuesto, para controlar mejor las correcciones que reciben.

distribuciones originales según las ECV, a posicionarse en torno al 110% de la media.

Cuadro 4.21.c
DISTRIBUCIÓN DE LA RENTA DISPONIBLE NETA PER CÁPITA EQUIVALENTE (2008)

	ECV b2004 (vhRentaAlapq2)	ECV b2013 (vhRentaAlapq2)	Corregida a partir de la b2004	Corregida a partir de la b2013
Nº de casos considerados	13.360	13.360	61.645	62.804
Hogares representados	17.279.282	17.279.282	17.324.391	17.324.391
Nº de personas representadas	45.870.296	45.870.296	45.983.169	45.983.169
Tamaño medio de los hogares	2,65	2,65	2,65	2,65
Mínimo	-75.001	-56.535	-710.497	-1.464.169
Percentil 5	4.694	5.777	4.822	5.364
Percentil 10	6.866	7.952	7.976	7.898
Percentil 25	10.558	11.693	12.231	11.916
Percentil 50	14.840	16.849	17.819	17.558
Percentil 75	20.831	23.594	25.388	25.502
Percentil 90	28.292	32.279	34.851	36.432
Percentil 95	33.973	39.430	43.046	45.133
Máximo	178.470	400.073	3.061.441	2.307.031
Media	16.571	19.124	20.554	20.618
Desviación Típica	9.616	11.993	24.263	19.643
Coeficiente de Variación de Pearson	58,0%	62,7%	118,0%	95,3%
Total agregado (Millones €)	760.130	877.232	945.157	948.095
Recursos acumulados hasta el Percentil 5	0,41%	0,83%	-0,32%	-0,43%
Recursos acumulados hasta el Percentil 10	2,19%	2,65%	1,29%	1,21%
Recursos acumulados hasta el Percentil 25	10,29%	10,54%	8,83%	8,56%
Recursos acumulados hasta el Percentil 50	29,51%	29,25%	27,00%	26,35%
Recursos acumulados hasta el Percentil 75	56,12%	55,30%	52,86%	51,93%
Recursos acumulados hasta el Percentil 90	77,83%	76,76%	74,28%	73,80%
Recursos acumulados hasta el Percentil 95	87,07%	86,04%	83,62%	83,52%
Índice de Gini	0,2992	0,3049	0,3455	0,3551
Umbral de Pobreza en la Clase de Referencia	8.904,22	10.109,33	10.691,45	10.534,55
Tasa de Pobreza	17,12%	17,33%	18,43%	18,90%
Brecha relativa de Pobreza de la UE	0,2872	0,2629	0,2849	0,2698

Fuente: Elaboración propia a partir de los datos de las ECV.

El desplazamiento de recursos hacia las zonas de mayores rentas, agudizan el efecto sobre los indicadores de desigualdad y pobreza. El índice de Gini pasa de estar en torno al 30% según las ECV, a valores de entorno al 35% según las distribuciones corregidas. El umbral de pobreza crece ligeramente con respecto de lo deducido directamente de las ECV para posicionarse en torno a los 10.600€, y las tasas de pobreza se incrementan aproximadamente en un punto y medio de porcentaje, manteniéndose la Brecha relativa de pobreza de la UE en niveles parecidos a los que estima directamente las ECV.

Pero, otra vez, si comparamos las dos últimas columnas de este cuadro, construidas para comprobar la robustez de las estimaciones que el método proporciona, observamos que las distribuciones obtenidas por ambas aplicaciones del método de corrección son bastante parecidas, con la misma excepción de sus valores más extremos, como ocurría en el cuadro 4.21.b. Las medias estimadas apenas se diferencian unas decenas de euros, los percentiles apenas difieren en unos pocos cientos de euros, apreciándose en este caso una mayor diferencia en el decil superior de la distribución con diferencias de unos 2.000 euros (que en todo caso supone una diferencia inferior al 5%); los recursos acumulados por tramos se muestran igualmente muy

parejos (con diferencias inferiores al medio punto porcentual); y los índices de Gini, umbrales de pobreza y tasa de pobreza presenta discrepancias realmente bajas.

5.3.2. Distribuciones personales la renta para el año 2011

Si observamos ahora los cuadros 4.22.a, 4.22.b y 4.22.c, análogos a los anteriores, pero contruidos para el año 2011, observamos características similares en el comportamiento del procedimiento de corrección propuesto.

Cuadro 4.22.a
**CONGRUENCIA DE LA DISTRIBUCIÓN PERSONAL DE LA RENTA
CORREGIDA CON CNE (2011)**

	BDH (b2004)	BDH (b2013)	CNE	BDH (b2004) Corregida	BDH (b2013) Corregida
Número de personas	46.354.779	46.354.779	46.736.256	46.736.257	46.736.257
Excedente Neto de Explotación	64.148.391.715	64.088.350.135	41.780.000.000	41.780.000.000	41.780.000.000
Renta mixta neta	43.172.335.963	61.102.168.874	86.437.000.000	86.437.000.000	86.437.000.000
Excedente de explotación neto / Renta mixta neta	107.320.727.678	125.190.519.009	128.217.000.000	128.217.000.000	128.217.000.000
Sueldos y Salarios	311.559.338.141	358.519.538.471	420.917.000.000	420.917.000.000	420.917.000.000
Cotizaciones sociales efectivas a cargo de empleadores	74.491.179.414	93.293.130.231	89.829.990.750	89.829.990.750	89.829.990.750
Cotizaciones sociales imputadas a cargo de empleadores			21.129.009.250	21.129.009.250	21.129.009.250
Cotizaciones sociales a cargo del empleador	74.491.179.414	93.293.130.231	110.959.000.000	110.959.000.000	110.959.000.000
Remuneración de Asalariados	386.050.517.554	451.812.668.701	531.876.000.000	531.876.000.000	531.876.000.000
Intereses (empleos)			21.278.000.000	21.278.000.000	21.278.000.000
Saldo (Rec-Emp) de Intereses, Dividendos y Ganacias de Inversiones de Capital			22.503.000.000	22.503.000.000	22.503.000.000
Rentas de la propiedad atribuidas a los asegurados (recursos)			16.211.000.000	16.211.000.000	16.211.000.000
Rentas de la tierra (rec-emp)	5.693.413.459	8.100.572.759	42.000.000	42.000.000	42.000.000
Rentas de la propiedad (rec-emp)	9.746.443.222	26.030.982.456	38.756.000.000	38.756.000.000	38.756.000.000
Saldo de rentas primarias netas	503.117.688.454	603.034.170.166	698.849.000.000	698.849.000.000	698.849.000.000
Impuestos sobre la renta	66.035.099.130	92.503.814.686	79.697.000.000	79.697.000.000	79.697.000.000
Otros impuestos corrientes	0	0	2.713.000.000	2.713.000.000	2.713.000.000
Impuestos corrientes sobre la renta, el patrimonio, etc.	66.035.099.130	92.503.814.686	82.410.000.000	82.410.000.000	82.410.000.000
Cotizaciones sociales (recursos)			187.000.000	187.000.000	187.000.000
Cotizaciones sociales (empleos)	74.491.179.414	93.293.130.231	153.825.000.000	153.825.000.000	153.825.000.000
Prest. Soc. distintas de transf. soc. en espec (Rec)	136.341.174.588	150.202.031.879	183.845.000.000	183.845.000.000	183.845.000.000
Prest. Soc. distintas de transf. soc. en espec (Emp)			171.000.000	171.000.000	171.000.000
Primas netas de seguro no vida			15.072.000.000	15.072.000.000	15.072.000.000
Indemnizaciones de seguro no vida (recursos)			14.818.000.000	14.818.000.000	14.818.000.000
Transferencias corrientes diversas (Recursos)	1.984.741.972	1.978.490.086	32.556.000.000	32.556.000.000	32.556.000.000
Transferencias corrientes diversas (empleos)	3.486.852.198	3.571.811.227	39.739.000.000	39.739.000.000	39.739.000.000
Otras transferencias corrientes (Recursos)	1.984.741.972	1.978.490.086	47.374.000.000	47.374.000.000	47.374.000.000
Otras transferencias corrientes (Empleos)	3.486.852.198	3.571.811.227	54.811.000.000	54.811.000.000	54.811.000.000
Renta disponible Neta	497.430.474.272	565.845.935.987	639.038.000.000	639.038.000.000	639.038.000.000
Consumo de Capital Fijo			43.602.000.000	43.602.000.000	43.602.000.000

Fuente: Elaboración propia a partir de los datos de las ECV y la CNE base 2010.

Nuevamente, los agregados para las distribuciones corregidas en este año, resultantes de la aplicación del procedimiento propuesto, tanto a los datos de la base 2004 como a los datos de la base 2013, respetan escrupulosamente los agregados marcados por CNE considerados como referencia.

Y la forma en que las estimaciones resultantes logran esta congruencia agregada con CNE, así como el grado de robustez de las estimaciones que el método propuesto proporciona, podemos estudiarla someramente en los cuadros 4.22.b y 4.23.c que presentamos a continuación.

En primer lugar, la renta disponible *per cápita* en los hogares en 2011, (cuadro 4.22.b), vuelve a mostrar correcciones muy acusadas entre el mínimo y el percentil 5, y entre el percentil 95 y el máximo, bastante mayores que en 2008, y que producen incrementos considerables en la dispersión de las distribuciones corregidas con respecto de las originales, pasando el Coeficiente de variación de Pearson a multiplicarse por 10.

También la distribución de recursos se ve desplazada especialmente hacia las rentas mayores que la mediana, agudizándose su efecto a partir del

percentil 95 de la distribución que pasa de acaparar aproximadamente un 14% de los recursos según las ECV, a aproximadamente un 16% después de la corrección.

Cuadro 4.22.b
DISTRIBUCIÓN DE LA RENTA DISPONIBLE NETA PER CÁPITA (2011)

	ECV b2004 (vhRentaAla pc)	ECV b2013 (vhRentaAla pc)	Corregida a partir de la b2004	Corregida a partir de la b2013
Nº de casos considerados	12.714	12.714	61.393	62.836
Hogares represenados	18.034.410	18.034.410	18.186.339	18.186.339
Nº de personas representadas	46.354.779	46.354.779	46.736.257	46.736.257
Tamaño medio de los hogares	2,57	2,57	2,57	2,57
Mínimo	-21.007	-21.007	-1.905.246	-1.980.339
Percentil 5	2.358	2.950	3.028	3.078
Percentil 10	3.591	4.257	4.566	4.668
Percentil 25	5.964	6.807	7.543	7.332
Percentil 50	9.436	10.668	11.536	11.378
Percentil 75	13.912	15.665	16.918	17.174
Percentil 90	19.670	22.061	24.209	24.315
Percentil 95	24.432	27.412	29.904	30.322
Máximo	122.327	196.633	27.940.381	63.920.032
Media	10.860	12.327	13.673	13.673
Desviación Típica	7.385	8.507	56.496	59.146
Coeficiente de Variación de Pearson	68,0%	69,0%	413,2%	432,6%
Total agregado (Millones €)	503.399	571.431	639.038	639.038
Recursos acumulados hasta el Percentil 5	0,48%	0,65%	0,02%	-0,15%
Recursos acumulados hasta el Percentil 10	1,85%	2,12%	1,42%	1,29%
Recursos acumulados hasta el Percentil 25	8,54%	8,99%	8,16%	7,98%
Recursos acumulados hasta el Percentil 50	26,22%	26,61%	25,59%	25,14%
Recursos acumulados hasta el Percentil 75	52,64%	52,79%	51,07%	50,73%
Recursos acumulados hasta el Percentil 90	75,23%	75,03%	72,95%	72,84%
Recursos acumulados hasta el Percentil 95	85,24%	84,92%	82,70%	82,69%
Índice de Gini	0,3458	0,3414	0,3663	0,3715

Fuente: Elaboración propia a partir de los datos de las ECV.

Y si comparamos las dos últimas columnas de este cuadro, construidas para comprobar la robustez de las estimaciones, observamos que las distribuciones obtenidas por ambas aplicaciones del método de corrección son muy parecidas, volviendo a coincidir sus medias y aproximándose sustancialmente tanto los percentiles de rentas como los de recursos, con la mencionada excepción de sus valores más extremos que parecen haber sido muy especialmente castigados por el procedimiento de corrección.

Características éstas que vuelven a reflejarse en el cuadro 4.22.c para la renta disponible equivalente per cápita en los hogares, en 2011. Nuevamente, con la excepción de las rentas extremas, las diferencias mostradas por ambas estimaciones realizadas son relativamente pequeñas. Las correcciones más acusadas se materializan nuevamente a partir de la mediana de la distribución, haciendo perder peso relativo a las clases bajas en el reparto total de recursos. Se observa, de forma análoga que en el caso anterior, un incremento importante en la dispersión de las distribuciones corregidas con respecto de las originales. Y de la misma forma, se observa que el desplazamiento de recursos hacia las zonas de mayores rentas, agudizan el efecto sobre los indicadores de desigualdad y pobreza, pasando el índice de

Gini de estar en torno al 31% según las ECV, a valores de entorno al 35% según las distribuciones corregidas.

El umbral de pobreza baja sólo unos 100€ con respecto del de 2008 (bajada menor que la que dicen las ECV); y las tasas de pobreza se incrementan aproximadamente en un punto y medio de porcentaje, estabilizándose también, en esta ocasión, la Brecha relativa de pobreza de la UE en torno al 28%.

Cuadro 4.22.c
DISTRIBUCIÓN DE LA RENTA DISPONIBLE NETA PER CÁPITA EQUIVALENTE (2011)

	ECV b2004 (vhRentaAlapq2)	ECV b2013 (vhRentaAlapq2)	Corregida a partir de la b2004	Corregida a partir de la b2013
Nº de casos considerados	12.714	12.714	61.393	62.836
Hogares representados	18.034.410	18.034.410	18.186.339	18.186.339
Nº de personas representadas	46.354.779	46.354.779	46.736.257	46.736.257
Tamaño medio de los hogares	2,57	2,57	2,57	2,57
Mínimo	-28.009	-28.009	-2.540.328	-2.640.451
Percentil 5	3.965	4.904	4.941	4.894
Percentil 10	5.973	7.131	7.529	7.597
Percentil 25	9.562	10.998	11.751	11.614
Percentil 50	14.226	16.030	17.694	17.358
Percentil 75	20.400	23.031	25.688	26.014
Percentil 90	28.039	31.371	35.554	35.715
Percentil 95	34.570	37.975	42.452	43.158
Máximo	135.817	196.633	53.731.501	133.166.734
Media	16.036	18.268	20.463	20.576
Desviación Típica	9.801	11.572	107.085	126.534
Coeficiente de Variación de Pearson	61,1%	63,3%	523,3%	615,0%
Total agregado (Millones €)	743.367	846.799	956.361	961.623
Recursos acumulados hasta el Percentil 5	0,57%	0,76%	0,19%	0,06%
Recursos acumulados hasta el Percentil 10	2,16%	2,43%	1,73%	1,62%
Recursos acumulados hasta el Percentil 25	9,60%	9,98%	8,91%	8,75%
Recursos acumulados hasta el Percentil 50	28,21%	28,43%	26,81%	26,28%
Recursos acumulados hasta el Percentil 75	54,76%	54,66%	52,85%	52,05%
Recursos acumulados hasta el Percentil 90	76,79%	76,59%	74,76%	74,17%
Recursos acumulados hasta el Percentil 95	86,43%	85,98%	84,21%	83,64%
Índice de Gini	0,3173	0,3151	0,3437	0,3525
Umbral de Pobreza en la Clase de Referencia	8.535,65	9.617,75	10.616,50	10.414,68
Tasa de Pobreza	19,70%	18,98%	20,45%	19,90%
Brecha relativa de Pobreza de la UE	0,3038	0,2760	0,2800	0,2729

Fuente: Elaboración propia a partir de los datos de las ECV.

Nuevamente, si comparamos las dos últimas columnas de este último cuadro, constatamos que las distribuciones obtenidas por ambas aplicaciones del método de corrección son bastante parecidas, con la mencionada excepción de sus valores más extremos. Las medias estimadas apenas se diferencian unas decenas de euros, los percentiles apenas difieren en unos pocos cientos de euros, apreciándose aún menor diferencia que entre las soluciones del año 2008; los recursos acumulados por tramos se muestran igualmente muy parejos (con diferencias inferiores al medio punto porcentual); y los índices de Gini, umbrales de pobreza, tasa de pobreza y Brecha relativa de pobreza de la UE presentan discrepancias realmente bajas.

En resumen, se observa una relativa robustez del método propuesto, propiciando estimaciones muy parecidas a partir de los dos conjuntos diferentes de datos de partida (base 2004 y base 2013), y encontrándose las nuevas distribuciones estimadas para ambas situaciones, cercanas a la distribución que presenta la nueva metodología de las ECV con base en los registros administrativos; si bien, corregida algo al alza por el efecto del descuadre aún existente con la CNE. Sólo la Brecha relativa de pobreza de la UE parece comportarse un poco más inestable, heredando las características del conjunto de datos empleado para la estimación en 2008; lo que sugiere la necesidad de profundizar en ello en un futuro.

5.4. Resultados comparativos con las ECV originales

Como se señala en Callealta y Romero (2013a y 2016), el uso de los microdatos que proporciona directamente las ECV podría introducir sesgos considerables en ciertos análisis. Así, dicen:

"... es evidente que las naturalezas de los ingresos de los hogares no son idénticas en toda la población, sino que se asocian fuertemente para diferentes grupos poblacionales con ciertas características comunes. Así, se precisa ser asalariado en algún momento para percibir rentas salariales al trabajo; o ser empresario o profesional libre para recibir rentas al trabajo por cuenta propia; o estar parado para percibir prestaciones sociales por desempleo; etc. Y, por tanto, niveles de ocultación en las declaraciones de determinados tipos de ingresos que pudieran ser muy diferentes, o desequilibrados, para los diferentes grupos poblacionales, podrían producir sesgos presumiblemente importantes en las mediciones de fenómenos sociales, como la desigualdad o la pobreza, mediante indicadores métricos objetivos. De este modo, y en caso extremo, si en una cierta subpoblación se ocultaran sistemáticamente la totalidad de sus ingresos reales, declarando ingresos nulos, todos los pertenecientes a dicha subpoblación se posicionarían con práctica seguridad bajo los umbrales de pobreza calculados para la población en su conjunto por métodos objetivos. Sin embargo, los integrantes de tal subpoblación no tienen por qué ser necesariamente pobres; simplemente han ocultado sus rentas. Pero no es éste el único efecto que se produciría; la ocultación de rentas de esta subpoblación sesgaría a la baja los indicadores de posición central de la distribución de ingresos (medias, mediana) y, en consecuencia, sesgará también a la baja el umbral de la pobreza, expulsando de la categoría de pobres a otros perceptores de rentas bajas, verdaderos pobres si declararon sus rentas reales."

El objetivo de este apartado final de esta Memoria, no es otro que el de presentar comparativamente las principales características de las distribuciones de las rentas personales derivadas para algunos de dichos colectivos de especial interés para el análisis económico, con el fin de observar comparativamente el efecto, positivo o negativo, que podrían estar induciendo las correcciones que proponemos sobre sus características. Y, para ello, iremos presentando en los cuadros siguientes, paralelamente para los

años 2008 y 2011 del estudio, las características de las distribuciones de la Renta Disponible Neta per cápita equivalente que para estos colectivos se derivan de:

- las ECV base 2004,
- las más actuales ECV base 2013 (que utilizan los registros administrativos),
- y las que presentan nuestras distribuciones corregidas a partir de las ECV base 2013, como información más actual.

En primer lugar, en el los cuadros 4.23.a y 4.23.b se compara, para 2008 y 2011, respectivamente, dichas distribuciones, para los colectivos de hogares formados en función del tipo de hábitat en que se ubican, de acuerdo con la definición que se recoge en las ECV, dependiendo del volumen de población.

Cuadro 4.23.a
**RENTA DISPONIBLE NETA PER CÁPITA EQUIVALENTE,
 SEGÚN TIPO DE HÁBITAT DEL HOGAR (2008)**

	España en su conjunto	Zona muy poblada	Zona medianamente poblada	Zona poco poblada
Resultados 2008 base 2004				
Número de Hogares representados en la muestra	17.279.282	9.086.661	3.632.129	4.560.493
% de hogares representados	100,00%	52,59%	21,02%	26,39%
Renta media	16.415	18.237	15.144	13.943
Índice de Posicionamiento de la Media (%)	100,00%	111,10%	92,26%	84,94%
Renta Total de la clase (Millones de euros)	752.938,6	429.716,3	153.843,0	169.379,3
Participación de la clase en la Renta Total (%)	100,00%	57,07%	20,43%	22,50%
Índice de Gini	0,3053	0,2989	0,3013	0,2932
Tasa de Pobreza (umbral de pobreza: 8.845€)	0,1800	0,1372	0,2097	0,2384
Brecha Relativa de Pobreza UE	0,2927	0,2965	0,3078	0,2654

	España en su conjunto	Zona muy poblada	Zona medianamente poblada	Zona poco poblada
Resultados 2008 base 2013				
Número de Hogares representados en la muestra	17.279.282	9.086.661	3.632.129	4.560.493
% de hogares representados	100,00%	52,59%	21,02%	26,39%
Renta media	18.982	21.038	17.296	16.403
Índice de Posicionamiento de la Media (%)	100,00%	110,83%	91,12%	86,42%
Renta Total de la clase (Millones de euros)	870.712,0	495.733,9	175.708,4	199.269,7
Participación de la clase en la Renta Total (%)	100,00%	56,93%	20,18%	22,89%
Índice de Gini	0,3102	0,3104	0,3067	0,2849
Tasa de Pobreza (umbral de pobreza: 10.024€)	0,1846	0,1504	0,2133	0,2268
Brecha Relativa de Pobreza UE	0,2581	0,2963	0,2852	0,1979

	España en su conjunto	Zona muy poblada	Zona medianamente poblada	Zona poco poblada
Resultados 2008 base 2013 Corregidos				
Número de Hogares representados en la muestra	17.324.391	9.112.584	3.637.707	4.574.101
% de hogares representados	100,00%	52,60%	21,00%	26,40%
Renta media	20.618	22.688	18.888	18.048
Índice de Posicionamiento de la Media (%)	100,00%	110,04%	91,61%	87,54%
Renta Total de la clase (Millones de euros)	948.095,4	536.088,0	192.126,4	219.881,0
Participación de la clase en la Renta Total (%)	100,00%	56,54%	20,26%	23,19%
Índice de Gini	0,3551	0,3512	0,3566	0,3412
Tasa de Pobreza (umbral de pobreza: 10.535€)	0,1890	0,1565	0,2161	0,2293
Brecha Relativa de Pobreza UE	0,2698	0,2722	0,2869	0,2565

Fuente: Elaboración propia a partir de los datos de las ECV.

Para el año 2008 se observa que, tras las correcciones que introducen tanto la base 2013 como el procedimiento de corrección propuesto, las posiciones relativas de las medias y las participaciones de las clases en la Renta Total en las *zonas muy pobladas* disminuyen muy ligeramente a favor de las *zonas poco pobladas*. Internamente, las tres clases aumentan armónicamente sus grados de desigualdad (muy significativamente) y las tasas de pobreza (muy ligeramente). Sin embargo, la Brecha relativa de pobreza de la UE baja en las *zonas muy pobladas*, al contrario de lo que ocurre en las *zonas poco pobladas*

en las que sube de forma considerable hasta 0,2565; casi hasta el nivel que presentaba según la base 2004 (quizás el descenso producido por la corrección de la base 2013, hasta el nivel 0,1979, era excesivo).

Cuadro 4.23.b
**RENTA DISPONIBLE NETA PER CÁPITA EQUIVALENTE,
 SEGÚN TIPO DE HÁBITAT DEL HOGAR (2011)**

	España en su conjunto	Zona muy poblada	Zona medianamente poblada	Zona poco poblada
Resultados 2011 base 2004				
Número de Hogares representados en la muestra	18.034.410	9.342.886	3.992.669	4.698.854
% de hogares representados	100,00%	51,81%	22,14%	26,05%
Renta media	15.867	17.539	15.062	13.400
Índice de Posicionamiento de la Media (%)	100,00%	110,54%	94,93%	84,45%
Renta Total de la clase (Millones de euros)	735.508,7	409.808,4	159.877,8	165.822,4
Participación de la clase en la Renta Total (%)	100,00%	55,72%	21,74%	22,55%
Índice de Gini	0,3231	0,3211	0,3115	0,3100
Tasa de Pobreza (umbral de pobreza: 8.445€)	0,2009	0,1770	0,2082	0,2398
Brecha Relativa de Pobreza UE	0,3224	0,3256	0,3325	0,3063

	España en su conjunto	Zona muy poblada	Zona medianamente poblada	Zona poco poblada
Resultados 2011 base 2013				
Número de Hogares representados en la muestra	18.034.410	9.342.886	3.992.669	4.698.854
% de hogares representados	100,00%	51,81%	22,14%	26,05%
Renta media	18.138	20.034	17.211	15.352
Índice de Posicionamiento de la Media (%)	100,00%	110,45%	94,89%	84,64%
Renta Total de la clase (Millones de euros)	840.767,5	468.090,3	182.694,7	189.982,5
Participación de la clase en la Renta Total (%)	100,00%	55,67%	21,73%	22,60%
Índice de Gini	0,3233	0,3252	0,3162	0,2983
Tasa de Pobreza (umbral de pobreza: 9.537€)	0,1980	0,1683	0,2141	0,2403
Brecha Relativa de Pobreza UE	0,3012	0,3410	0,3018	0,2510

	España en su conjunto	Zona muy poblada	Zona medianamente poblada	Zona poco poblada
Resultados 2011 base 2013 Corregidos				
Número de Hogares representados en la muestra	18.186.339	9.417.800	4.028.586	4.739.953
% de hogares representados	100,00%	51,79%	22,15%	26,06%
Renta media	20.576	22.109	19.591	18.528
Índice de Posicionamiento de la Media (%)	100,00%	107,45%	95,21%	90,05%
Renta Total de la clase (Millones de euros)	961.623,0	520.625,7	209.780,3	231.217,0
Participación de la clase en la Renta Total (%)	100,00%	54,14%	21,82%	24,04%
Índice de Gini	0,3525	0,3484	0,3474	0,3498
Tasa de Pobreza (umbral de pobreza: 10.415€)	0,1990	0,1784	0,2175	0,2218
Brecha Relativa de Pobreza UE	0,2729	0,2941	0,2721	0,2385

Fuente: Elaboración propia a partir de los datos de las ECV.

En el año 2011, las posiciones relativas de las medias y las participaciones de las clases en la Renta Total en las zonas *muy pobladas* disminuyen algo más acentuadamente, igual que en 2008, a favor de las zonas *poco pobladas*. Internamente, las tres clases aumentan armónicamente sus grados de desigualdad (más en las zonas *poco pobladas*, que la base 2013 había corregido a la baja). La tasa de pobreza aumenta en las zonas *muy pobladas* y disminuye más en las zonas *poco pobladas*, mientras que las Brechas relativas de pobreza UE bajan en todas las zonas, pero mantienen sus posiciones relativas presentando Brechas mayores cuanto más poblada es la zona.

En los cuadros 4.24.a y 4.24.b podemos ver el efecto comparativo de las correcciones realizadas en 2008 y 2011, respectivamente, sobre la Renta Disponible Neta per cápita equivalente de los hogares, según la categoría socio-económica del sustentador principal del hogar.

Observamos en ambos cómo, de forma general, la corrección que introduce la base 2013 produce un incremento general de rentas en todas las categorías, muy especialmente en las *empresarios agrarios y no agrarios*, con y sin asalariados, y en la de *cuadros superiores no agrarios*. Pero, además, las correcciones no son proporcionales en todas las categorías, produciéndose cambios sustanciales en el posicionamiento de éstas con respecto de la media nacional. En este sentido, al aplicar el método propuesto, éstas correcciones se ven matizadas en intensidad, pero manteniendo prácticamente su posicionamiento ordinal; con las excepciones de los *empresarios agrarios sin asalariados* y *otros asalariados no clasificados* que ven aumentadas algo más sus rentas en 2011.

Cuadro 4.24.a

RENTA DISPONIBLE NETA PER CÁPITA EQUIVALENTE, SEGÚN CATEGORÍA SOCIO-ECONÓMICA DEL SUSTENTADOR PRINCIPAL DEL HOGAR (2008)

	España en su conjunto	Empres. Agrar. con asal	Empres. No Agr. con asal	Empres. Agrar. sin asal	Empres. No Agr. sin asal	Asalariados Agrarios	Cuadros Super. No Agr.	Resto Empleados cualificad. No Agr.	Empleados no cualif. No Agrar.	Fuerzas Armadas	Otros Asalariados no clasific.	Parados	Formación	Retirados	Otros Inactivos	Otros Inactivos
Resultados 2008 base 2004																
Número de Hogares representados en la muestra	17.279.282	39.455	522.134	108.252	775.146	181.282	1.666.320	5.298.978	1.033.020	68.154	267.894	1.573.323	31.092	4.182.566	1.509.213	1.540.305
% de hogares representados	100,00%	0,23%	3,02%	0,63%	4,49%	1,05%	9,64%	30,67%	5,96%	0,39%	1,55%	9,11%	0,18%	24,21%	8,73%	8,91%
Renta media	16.415	13.615	16.112	11.922	13.729	11.812	27.327	17.241	13.191	19.883	14.318	10.879	11.479	15.993	12.494	12.494
Índice de Posicionamiento de la Media (%)	100,00%	82,94%	98,16%	72,63%	83,64%	71,96%	166,48%	105,04%	80,36%	121,13%	87,23%	66,28%	69,93%	97,43%	76,12%	76,12%
Renta Total de la clase (Millones de euros)	752.938,6	1.853,7	25.782,9	4.231,8	31.033,5	7.430,6	128.278,3	262.864,3	41.548,1	4.087,9	14.657,4	50.747,8	655,6	145.995,3	33.560,7	34.216
Participación de la clase en la Renta Total (%)	100,00%	0,25%	3,42%	0,56%	4,12%	0,99%	17,04%	34,91%	5,52%	0,54%	1,95%	6,74%	0,09%	19,39%	4,46%	4,54%
Índice de Gini	0,3053	0,3064	0,4218	0,3466	0,4108	0,2853	0,2191	0,2500	0,2547	0,1880	0,2185	0,3674	0,3596	0,2523	0,3118	0,3118
Tasa de Pobreza (umbral de pobreza: 8.845€)	0,1800	0,3865	0,2938	0,3225	0,3273	0,3411	0,0158	0,1104	0,2443	0,0402	0,1539	0,4318	0,4455	0,1215	0,2972	0,2972
Brecha Relativa de Pobreza UE	0,2927	0,1061	0,3592	0,4388	0,4689	0,3026	0,3568	0,2354	0,2746	0,5728	0,0681	0,3753	0,5461	0,1624	0,3827	0,3827
	España en su conjunto	Empres. Agrar. con asal	Empres. No Agr. con asal	Empres. Agrar. sin asal	Empres. No Agr. sin asal	Asalariados Agrarios	Cuadros Super. No Agr.	Resto Empleados cualificad. No Agr.	Empleados no cualif. No Agrar.	Fuerzas Armadas	Otros Asalariados no clasific.	Parados	Formación	Retirados	Otros Inactivos	Otros Inactivos
Resultados 2008 base 2013																
Número de Hogares representados en la muestra	17.279.282	45.808	554.939	120.542	830.061	167.601	1.629.146	5.221.635	1.016.875	72.753	251.714	1.631.807	31.341	4.140.688	1.541.918	1.573.259
% de hogares representados	100,00%	0,27%	3,21%	0,70%	4,80%	0,97%	9,43%	30,22%	5,88%	0,42%	1,46%	9,44%	0,18%	23,96%	8,92%	9,10%
Renta media	18.982	20.193	25.989	18.293	17.840	12.204	31.223	19.190	14.129	22.363	17.021	12.368	10.498	18.623	14.857	14.857
Índice de Posicionamiento de la Media (%)	100,00%	106,38%	136,91%	96,37%	93,98%	64,29%	164,49%	101,09%	74,43%	117,81%	89,47%	65,19%	55,31%	98,11%	78,27%	78,27%
Renta Total de la clase (Millones de euros)	870.712,0	3.312,1	44.039,6	7.280,5	44.361,2	6.987,9	140.993,6	289.534,8	42.959,7	4.812,9	16.062,3	60.308,6	592,4	167.433,0	41.516,9	42.109
Participación de la clase en la Renta Total (%)	100,00%	0,38%	5,06%	0,84%	5,09%	0,80%	16,19%	33,25%	4,93%	0,55%	1,84%	6,93%	0,07%	19,23%	4,77%	4,84%
Índice de Gini	0,3102	0,2939	0,3634	0,3172	0,3473	0,2631	0,2301	0,2610	0,2473	0,1792	0,2253	0,3752	0,3151	0,2594	0,3011	0,3011
Tasa de Pobreza (umbral de pobreza: 10.024€)	0,1846	0,0874	0,1550	0,1946	0,2489	0,4421	0,0208	0,1298	0,2845	0,0060	0,1661	0,4606	0,4573	0,1238	0,2785	0,2785
Brecha Relativa de Pobreza UE	0,2581	0,2588	0,3009	0,2024	0,3151	0,2230	0,3639	0,2075	0,2640	0,1519	0,0931	0,3706	0,5385	0,1498	0,2749	0,2749
	España en su conjunto	Empres. Agrar. con asal	Empres. No Agr. con asal	Empres. Agrar. sin asal	Empres. No Agr. sin asal	Asalariados Agrarios	Cuadros Super. No Agr.	Resto Empleados cualificad. No Agr.	Empleados no cualif. No Agrar.	Fuerzas Armadas	Otros Asalariados no clasific.	Parados	Formación	Retirados	Otros Inactivos	Otros Inactivos
Resultados 2008 base 2013 Corregidos																
Número de Hogares representados en la muestra	17.324.391	45.927	556.285	121.138	832.366	167.685	1.633.673	5.235.995	1.018.863	72.788	251.757	1.634.958	31.392	4.152.779	1.546.351	1.577.743
% de hogares representados	100,00%	0,27%	3,21%	0,70%	4,80%	0,97%	9,43%	30,22%	5,88%	0,42%	1,45%	9,44%	0,18%	23,97%	8,93%	9,11%
Renta media	20.618	20.442	26.981	19.095	20.981	14.883	33.665	21.366	16.323	23.761	17.100	15.652	10.342	17.858	16.103	16.103
Índice de Posicionamiento de la Media (%)	100,00%	99,14%	130,86%	92,61%	101,76%	72,18%	163,28%	103,63%	79,17%	115,24%	82,94%	75,91%	50,16%	86,61%	78,10%	78,10%
Renta Total de la clase (Millones de euros)	948.095,4	3.362,2	45.824,4	7.634,1	52.310,5	8.526,0	152.449,6	323.239,2	49.718,5	5.117,6	16.141,3	76.462,4	583,9	160.987,5	45.117,0	45.701
Participación de la clase en la Renta Total (%)	100,00%	0,35%	4,83%	0,81%	5,52%	0,90%	16,08%	34,09%	5,24%	0,54%	1,70%	8,06%	0,06%	16,98%	4,76%	4,82%
Índice de Gini	0,3551	0,2983	0,3735	0,3367	0,4138	0,3048	0,2600	0,2971	0,3069	0,1949	0,2279	0,4130	0,3741	0,3456	0,4155	0,4155
Tasa de Pobreza (umbral de pobreza: 10.535€)	0,1890	0,1076	0,1383	0,1916	0,2142	0,3085	0,0362	0,1160	0,2450	0,0133	0,1207	0,3594	0,4724	0,2312	0,3554	0,3554
Brecha Relativa de Pobreza UE	0,2698	0,1622	0,2822	0,2355	0,3317	0,1795	0,2987	0,2219	0,1976	0,9134	0,1654	0,3524	0,5925	0,2636	0,2661	0,2661

Fuente: Elaboración propia a partir de los datos de las ECV.

Cuadro 4.24.b

RENTA DISPONIBLE NETA PER CÁPITA EQUIVALENTE, SEGÚN CATEGORÍA SOCIO-ECONÓMICA DEL SUSTENTADOR PRINCIPAL DEL HOGAR (2011)

	España en su conjunto	Empres. Agrar. con asal	Empres. No Agr. con asal	Empres. Agrar. sin asal	Empres. No Agr. sin asal	Asalariados Agrarios	Cuadros Super. No Agr.	Resto Empleados cualificad. No Agr.	Empleados no cualif. No Agrar.	Fuerzas Armadas	Otros Asalariados no clasific.	Parados	Formación	Retirados	Otros Inactivos	Otros Inactivos
Resultados 2011 base 2004																
Número de Hogares representados en la muestra	18.034.410	63.506	476.749	101.358	699.018	267.714	1.779.121	5.358.738	730.708	70.586	162.620	2.041.611	43.767	4.494.791	1.742.426	1.786.193
% de hogares representados	100,00%	0,35%	2,64%	0,56%	3,88%	1,48%	9,87%	29,71%	4,05%	0,39%	0,90%	11,32%	0,24%	24,92%	9,66%	9,90%
Renta media	15.867	16.250	18.755	14.324	15.742	11.122	25.972	16.051	11.530	19.918	15.725	9.062	4.453	16.594	12.421	12.421
Índice de Posicionamiento de la Media (%)	100,00%	102,42%	118,20%	90,28%	99,21%	70,09%	163,69%	101,16%	72,67%	125,53%	99,11%	57,11%	28,07%	104,58%	78,28%	78,28%
Renta Total de la clase (Millones de euros)	735.508,7	3.222,0	25.874,6	4.106,6	31.323,4	9.464,3	126.299,2	247.685,9	25.028,0	4.301,7	8.518,1	50.328,9	0,0	158.902,4	40.147,3	40.147
Participación de la clase en la Renta Total (%)	100,00%	0,44%	3,52%	0,56%	4,26%	1,29%	17,17%	33,68%	3,40%	0,58%	1,16%	6,84%	0,00%	21,60%	5,46%	5,46%
Índice de Gini	0,3231	0,2754	0,4486	0,4465	0,4877	0,2823	0,2223	0,2532	0,2653	0,2045	0,2324	0,4016	0,0000	0,2551	0,2972	0,2972
Tasa de Pobreza (umbral de pobreza: 8.445€)	0,2009	0,1216	0,3176	0,4153	0,3927	0,3529	0,0163	0,1337	0,2993	0,0071	0,1104	0,5353	0,0000	0,0920	0,2733	0,2733
Brecha Relativa de Pobreza UE	0,3224	0,3251	0,4219	0,2715	0,4114	0,2796	0,4819	0,2354	0,2489	0,0416	0,3007	0,4433	0,0000	0,1548	0,3222	0,3222
Resultados 2011 base 2013																
Número de Hogares representados en la muestra	18.034.410	68.585	542.442	102.244	760.249	250.673	1.755.374	5.225.143	735.268	70.655	168.450	2.111.085	52.041	4.411.346	1.777.537	1.829.578
% de hogares representados	100,00%	0,38%	3,01%	0,57%	4,22%	1,39%	9,73%	28,97%	4,08%	0,39%	0,93%	11,71%	0,29%	24,46%	9,86%	10,14%
Renta media	18.138	16.866	25.029	15.455	16.986	11.425	30.591	18.032	12.804	21.008	18.036	11.498	10.102	18.518	14.117	14.117
Índice de Posicionamiento de la Media (%)	100,00%	92,99%	137,99%	85,21%	93,65%	64,09%	168,66%	99,42%	70,59%	115,83%	99,44%	63,39%	55,70%	102,10%	77,83%	77,83%
Renta Total de la clase (Millones de euros)	840.767,5	3.581,4	39.772,7	4.700,8	37.986,0	9.274,7	146.102,2	269.648,0	28.367,1	4.550,6	9.954,1	66.401,2	0,0	172.055,6	47.227,3	47.227
Participación de la clase en la Renta Total (%)	100,00%	0,43%	4,73%	0,56%	4,52%	1,10%	17,38%	32,07%	3,37%	0,54%	1,18%	7,90%	0,00%	20,46%	5,62%	5,62%
Índice de Gini	0,3233	0,2670	0,3534	0,3145	0,4015	0,2815	0,2421	0,2557	0,2740	0,2170	0,2460	0,4173	0,0000	0,2604	0,3018	0,3018
Tasa de Pobreza (umbral de pobreza: 9.537€)	0,1980	0,1474	0,1381	0,2371	0,3236	0,4739	0,0095	0,1355	0,3225	0,0069	0,1382	0,5030	0,0000	0,1007	0,2996	0,2996
Brecha Relativa de Pobreza UE	0,3012	0,1025	0,2826	0,3544	0,3248	0,2091	0,0978	0,2306	0,2983	0,0117	0,1778	0,4176	0,0000	0,1958	0,2807	0,2807
Resultados 2011 base 2013 Corregidos																
Número de Hogares representados en la muestra	18.186.339	68.917	546.900	103.512	766.257	252.308	1.769.783	5.269.979	741.133	71.183	169.784	2.126.737	52.420	4.451.627	1.792.419	1.844.839
% de hogares representados	100,00%	0,38%	3,01%	0,57%	4,21%	1,39%	9,73%	28,98%	4,08%	0,39%	0,93%	11,69%	0,29%	24,48%	9,86%	10,14%
Renta media	20.576	18.315	26.754	22.435	18.183	14.047	35.813	20.612	15.244	23.526	25.566	14.813	14.166	19.266	15.039	15.039
Índice de Posicionamiento de la Media (%)	100,00%	89,01%	130,03%	109,04%	88,37%	68,27%	174,06%	100,18%	74,09%	114,34%	124,25%	71,99%	68,85%	93,63%	73,09%	73,09%
Renta Total de la clase (Millones de euros)	961.623,0	3.908,1	42.863,3	6.897,9	40.986,5	11.275,3	172.425,8	310.814,7	34.033,1	5.132,5	14.222,9	86.192,3	0,0	180.625,1	50.706,9	50.707
Participación de la clase en la Renta Total (%)	100,00%	0,41%	4,46%	0,72%	4,26%	1,17%	17,93%	32,32%	3,54%	0,53%	1,48%	8,96%	0,00%	18,78%	5,27%	5,27%
Índice de Gini	0,3525	0,3045	0,3469	0,4846	0,4015	0,3200	0,2737	0,2782	0,2923	0,2281	0,4565	0,4343	0,0000	0,3118	0,3568	0,3568
Tasa de Pobreza (umbral de pobreza: 10.415€)	0,1990	0,1559	0,1481	0,2221	0,3104	0,3645	0,0125	0,1321	0,2788	0,0427	0,1787	0,4201	0,0000	0,1670	0,3356	0,3356
Brecha Relativa de Pobreza UE	0,2729	0,3077	0,2989	0,2798	0,3163	0,2252	0,2654	0,1897	0,2874	0,0549	0,2384	0,3742	0,0000	0,1865	0,2865	0,2865

Fuente: Elaboración propia a partir de los datos de las ECV.

En dichos cuadros podemos ver también que *parados* y *retirados* sufren ligeras correcciones contrarias al aplicar el método propuesto con respecto de la corrección realizada en la base 2013. Así, los *parados* ven incrementarse algo más sus rentas, mientras que los *retirados* las ven reducirse un poco.

Por su parte, se observan dos categorías en las que la pobreza prácticamente no existe y en las que la desigualdad es relativamente baja: *cuadros superiores no agrarios* y *fuerzas armadas*. Y las mismas correcciones de renta que introdujeron las ECV base 2013, afectaron heterogéneamente a las tasas de pobreza de los grupos más afectados antes señalados, propiciando tasas intuitivamente más aceptables. La nueva metodología respeta bastante estas cifras, matizándolas ligeramente según el año. Sólo se aprecian cambios algo más importantes y del mismo signo en la categoría de *asalariados agrarios*, en la que la tasa de pobreza disminuye de forma sustancial; en la categoría de *parados*, que disminuyen su tasa de pobreza; y en la de *retirados*, que la ven aumentar.

Se observa en general, que las nuevas correcciones producen un incremento en torno a tres puntos porcentuales en el índice de desigualdad de Gini, que se transmite al alza a prácticamente todas las categorías, aunque no de forma homogénea.

En el cuadro 4.25.a y el cuadro 4.25.b podemos ver el efecto comparativo de las correcciones realizadas en 2008 y 2011, respectivamente, sobre la Renta Disponible Neta per cápita equivalente de los hogares, según el grupo de edad al que pertenece el sustentador principal del hogar.

Cuadro 4.25.a
**RENTA DISPONIBLE NETA PER CÁPITA EQUIVALENTE,
 SEGÚN GRUPO DE EDAD DEL SUSTENTADOR PRINCIPAL DEL HOGAR (2008)**

	España en su conjunto	más de 16 años y hasta 25 años	más de 25 años y hasta 35 años	más de 35 años y hasta 45 años	más de 45 años y hasta 55 años	más de 55 años y hasta 65 años	más de 65 años
Resultados 2008 base 2004							
Número de Hogares representados en la muestra	17.279.282	436.728	3.106.855	3.991.577	3.059.186	2.567.720	4.117.216
% de hogares representados	100,00%	2,53%	17,98%	23,10%	17,70%	14,86%	23,83%
Renta media	16.415	11.692	16.807	15.910	17.278	18.324	14.804
Índice de Posicionamiento de la Media (%)	100,00%	71,23%	102,39%	96,93%	105,26%	111,63%	90,19%
Renta Total de la clase (Millones de euros)	752.938,6	13.236,3	134.869,3	197.125,9	169.063,1	121.900,7	116.743,3
Participación de la clase en la Renta Total (%)	100,00%	1,76%	17,91%	26,18%	22,45%	16,19%	15,51%
Índice de Gini	0,3053	0,3208	0,3133	0,3398	0,2874	0,2920	0,2451
Tasa de Pobreza (umbral de pobreza: 8.845€)	0,1800	0,3622	0,2011	0,2228	0,1544	0,1194	0,1481
Brecha Relativa de Pobreza UE	0,2927	0,4126	0,2947	0,3315	0,3026	0,3043	0,1715

	España en su conjunto	más de 16 años y hasta 25 años	más de 25 años y hasta 35 años	más de 35 años y hasta 45 años	más de 45 años y hasta 55 años	más de 55 años y hasta 65 años	más de 65 años
Resultados 2008 base 2013							
Número de Hogares representados en la muestra	17.279.282	404.206	3.046.454	4.017.637	3.117.960	2.601.672	4.091.354
% de hogares representados	100,00%	2,34%	17,63%	23,25%	18,04%	15,06%	23,68%
Renta media	18.982	12.535	17.711	18.277	20.718	22.021	17.296
Índice de Posicionamiento de la Media (%)	100,00%	66,03%	93,30%	96,29%	109,15%	116,01%	91,12%
Renta Total de la clase (Millones de euros)	870.712,0	12.024,1	138.604,4	227.422,8	208.171,7	149.445,7	135.043,2
Participación de la clase en la Renta Total (%)	100,00%	1,38%	15,92%	26,12%	23,91%	17,16%	15,51%
Índice de Gini	0,3102	0,3145	0,3098	0,3396	0,2938	0,3001	0,2597
Tasa de Pobreza (umbral de pobreza: 10.024€)	0,1846	0,3874	0,2277	0,2351	0,1422	0,1109	0,1544
Brecha Relativa de Pobreza UE	0,2581	0,3528	0,2992	0,2893	0,2556	0,2683	0,1607

	España en su conjunto	más de 16 años y hasta 25 años	más de 25 años y hasta 35 años	más de 35 años y hasta 45 años	más de 45 años y hasta 55 años	más de 55 años y hasta 65 años	más de 65 años
Resultados 2008 base 2013 Corregidos							
Número de Hogares representados en la muestra	17.324.391	405.031	3.052.963	4.027.504	3.126.019	2.608.809	4.104.066
% de hogares representados	100,00%	2,34%	17,62%	23,25%	18,04%	15,06%	23,69%
Renta media	20.618	15.483	20.196	20.715	22.845	23.777	15.908
Índice de Posicionamiento de la Media (%)	100,00%	75,09%	97,95%	100,47%	110,80%	115,32%	77,15%
Renta Total de la clase (Millones de euros)	948.095,4	14.876,0	158.370,3	258.367,0	230.110,1	161.805,5	124.566,5
Participación de la clase en la Renta Total (%)	100,00%	1,57%	16,70%	27,25%	24,27%	17,07%	13,14%
Índice de Gini	0,3551	0,4174	0,3460	0,3559	0,3368	0,3415	0,3583
Tasa de Pobreza (umbral de pobreza: 10.535€)	0,1890	0,3460	0,1935	0,1825	0,1401	0,1296	0,2898
Brecha Relativa de Pobreza UE	0,2698	0,3922	0,2876	0,2691	0,2533	0,2620	0,2633

Fuente: Elaboración propia a partir de los datos de las ECV.

Con respecto de las rentas medias en estas categorías vemos que, a pesar de las correcciones al alza que produce la base 2013, éstas respetan bastante las posiciones relativas de las categorías con respecto de la media nacional, con la excepción de la categoría de hogares con sustentador principal entre 25 y 35 años, que corrige relativamente a la baja con respecto del resto de categorías, llevándola a situarse un poco por encima del 90% de la media nacional. Y las nuevas correcciones que proponemos respetan las de la base 2013, con la excepción de la categoría de hogares con sustentadores principales de 65 o más años, donde se ubican fundamentalmente los hogares con sustentadores principales en la categoría socioeconómica de *retirados*, que ve descender su posición relativa a posiciones del 77,15% en 2008 y del 84,78% en 2011.

Cuadro 4.25.b
**RENTA DISPONIBLE NETA PER CÁPITA EQUIVALENTE,
 SEGÚN GRUPO DE EDAD DEL SUSTENTADOR PRINCIPAL DEL HOGAR (2011)**

	España en su conjunto	más de 16 años y hasta 25 años	más de 25 años y hasta 35 años	más de 35 años y hasta 45 años	más de 45 años y hasta 55 años	más de 55 años y hasta 65 años	más de 65 años
Resultados 2011 base 2004							
Número de Hogares representados en la muestra	18.034.410	275.743	2.815.607	4.169.608	3.444.739	2.778.943	4.549.769
% de hogares representados	100,00%	1,53%	15,61%	23,12%	19,10%	15,41%	25,23%
Renta media	15.867	9.798	15.560	15.263	16.168	18.017	15.414
Índice de Posicionamiento de la Media (%)	100,00%	61,75%	98,07%	96,20%	101,90%	113,55%	97,14%
Renta Total de la clase (Millones de euros)	735.508,7	7.015,2	107.049,1	191.914,7	170.443,8	124.332,8	134.753,0
Participación de la clase en la Renta Total (%)	100,00%	0,95%	14,55%	26,09%	23,17%	16,90%	18,32%
Índice de Gini	0,3231	0,3759	0,3455	0,3440	0,3289	0,3139	0,2513
Tasa de Pobreza (umbral de pobreza: 8.445€)	0,2009	0,4508	0,2446	0,2473	0,2127	0,1445	0,1095
Brecha Relativa de Pobreza UE	0,3224	0,4760	0,3324	0,3515	0,3179	0,2945	0,1821

	España en su conjunto	más de 16 años y hasta 25 años	más de 25 años y hasta 35 años	más de 35 años y hasta 45 años	más de 45 años y hasta 55 años	más de 55 años y hasta 65 años	más de 65 años
Resultados 2011 base 2013							
Número de Hogares representados en la muestra	18.034.410	270.642	2.848.508	4.144.586	3.484.729	2.795.486	4.490.459
% de hogares representados	100,00%	1,50%	15,79%	22,98%	19,32%	15,50%	24,90%
Renta media	18.138	11.421	16.620	17.729	18.591	21.257	17.398
Índice de Posicionamiento de la Media (%)	100,00%	62,97%	91,64%	97,75%	102,50%	117,20%	95,92%
Renta Total de la clase (Millones de euros)	840.767,5	7.718,4	116.181,3	221.138,4	199.674,4	147.689,4	148.365,6
Participación de la clase en la Renta Total (%)	100,00%	0,92%	13,82%	26,30%	23,75%	17,57%	17,65%
Índice de Gini	0,3233	0,3524	0,3209	0,3496	0,3223	0,3311	0,2524
Tasa de Pobreza (umbral de pobreza: 9.537€)	0,1980	0,4422	0,2328	0,2512	0,1963	0,1503	0,1134
Brecha Relativa de Pobreza UE	0,3012	0,3987	0,3690	0,3091	0,2702	0,2811	0,1730

	España en su conjunto	más de 16 años y hasta 25 años	más de 25 años y hasta 35 años	más de 35 años y hasta 45 años	más de 45 años y hasta 55 años	más de 55 años y hasta 65 años	más de 65 años
Resultados 2011 base 2013 Corregidos							
Número de Hogares representados en la muestra	18.186.339	272.637	2.871.401	4.177.132	3.514.311	2.819.870	4.530.988
% de hogares representados	100,00%	1,50%	15,79%	22,97%	19,32%	15,51%	24,91%
Renta media	20.576	13.595	19.286	20.515	22.494	23.538	17.444
Índice de Posicionamiento de la Media (%)	100,00%	66,07%	93,74%	99,70%	109,33%	114,40%	84,78%
Renta Total de la clase (Millones de euros)	961.623,0	9.249,7	135.880,3	257.890,1	243.605,0	164.919,3	150.078,6
Participación de la clase en la Renta Total (%)	100,00%	0,96%	14,13%	26,82%	25,33%	17,15%	15,61%
Índice de Gini	0,3525	0,3407	0,3284	0,3617	0,3689	0,3602	0,2950
Tasa de Pobreza (umbral de pobreza: 10.415€)	0,1990	0,3815	0,2069	0,2254	0,1807	0,1557	0,1976
Brecha Relativa de Pobreza UE	0,2729	0,4073	0,3093	0,2975	0,2786	0,2820	0,1879

Fuente: Elaboración propia a partir de los datos de las ECV.

La corrección de la base 2013 no introdujo grandes cambios ni en el índice de desigualdad de Gini, ni en la tasa de pobreza, pero aminoró la Brecha de pobreza en los hogares en que su sustentador principal tenía *hasta 25 años o entre 35 y 55 años*.

Por su parte, las nuevas correcciones incrementan la desigualdad de forma general y hacen disminuir las tasas de pobreza para los grupos de edad propios de la actividad económica (*hasta 65 años*), tanto más cuanto más joven es el sustentador principal del hogar, haciéndola crecer en el grupo de *mayores de 65* (presumiblemente inactivos). Sin embargo, también hacen aumentar las Brechas de pobreza, tanto más cuanto más joven es el sustentador principal del hogar, haciéndola crecer especialmente en los dos grupos de sustentadores principales más jóvenes.

En el cuadro 4.26.a y el cuadro 4.26.b podemos ver el efecto comparativo de las correcciones realizadas en 2008 y 2001 respectivamente, sobre la Renta Disponible Neta per cápita equivalente de los hogares, según el sexo del sustentador principal del hogar.

Cuadro 4.26.a
**RENTA DISPONIBLE NETA PER CÁPITA EQUIVALENTE,
SEGÚN GÉNERO DEL SUSTENTADOR PRINCIPAL DEL HOGAR (2008)**

	España en su conjunto	Hombre	Mujer
Resultados 2008 base 2004			
Número de Hogares representados en la muestra	17.279.282	11.812.863	5.466.419
% de hogares representados	100,00%	68,36%	31,64%
Renta media	16.415	16.445	16.329
Índice de Posicionamiento de la Media (%)	100,00%	100,19%	99,48%
Renta Total de la clase (Millones de euros)	752.938,6	553.208,0	199.730,6
Participación de la clase en la Renta Total (%)	100,00%	73,47%	26,53%
Índice de Gini	0,3053	0,2973	0,3269
Tasa de Pobreza (umbral de pobreza: 8.845€)	0,1800	0,1734	0,1983
Brecha Relativa de Pobreza UE	0,2927	0,2746	0,3546

	España en su conjunto	Hombre	Mujer
Resultados 2008 base 2013			
Número de Hogares representados en la muestra	17.279.282	11.703.937	5.575.345
% de hogares representados	100,00%	67,73%	32,27%
Renta media	18.982	19.050	18.801
Índice de Posicionamiento de la Media (%)	100,00%	100,36%	99,05%
Renta Total de la clase (Millones de euros)	870.712,0	635.863,3	234.848,6
Participación de la clase en la Renta Total (%)	100,00%	73,03%	26,97%
Índice de Gini	0,3102	0,3046	0,3250
Tasa de Pobreza (umbral de pobreza: 10.024€)	0,1846	0,1760	0,2075
Brecha Relativa de Pobreza UE	0,2581	0,2486	0,2923

	España en su conjunto	Hombre	Mujer
Resultados 2008 base 2013 Corregidos			
Número de Hogares representados en la muestra	17.324.391	11.734.695	5.589.696
% de hogares representados	100,00%	67,74%	32,26%
Renta media	20.618	20.573	20.740
Índice de Posicionamiento de la Media (%)	100,00%	99,78%	100,59%
Renta Total de la clase (Millones de euros)	948.095,4	688.419,3	259.676,0
Participación de la clase en la Renta Total (%)	100,00%	72,61%	27,39%
Índice de Gini	0,3551	0,3420	0,3884
Tasa de Pobreza (umbral de pobreza: 10.535€)	0,1890	0,1771	0,2207
Brecha Relativa de Pobreza UE	0,2698	0,2528	0,3106

Fuente: Elaboración propia a partir de los datos de las ECV.

Cuadro 4.26.b
**RENTA DISPONIBLE NETA PER CÁPITA EQUIVALENTE,
 SEGÚN GÉNERO DEL SUSTENTADOR PRINCIPAL DEL HOGAR (2011)**

	España en su conjunto	Hombre	Mujer
Resultados 2011 base 2004			
Número de Hogares representados en la muestra	18.034.410	11.788.949	6.245.461
% de hogares representados	100,00%	65,37%	34,63%
Renta media	15.867	15.937	15.703
Índice de Posicionamiento de la Media (%)	100,00%	100,44%	98,97%
Renta Total de la clase (Millones de euros)	735.508,7	517.460,0	218.048,7
Participación de la clase en la Renta Total (%)	100,00%	70,35%	29,65%
Índice de Gini	0,3231	0,3151	0,3415
Tasa de Pobreza (umbral de pobreza: 8.445€)	0,2009	0,1901	0,2260
Brecha Relativa de Pobreza UE	0,3224	0,3135	0,3472

	España en su conjunto	Hombre	Mujer
Resultados 2011 base 2013			
Número de Hogares representados en la muestra	18.034.410	11.752.151	6.282.259
% de hogares representados	100,00%	65,17%	34,83%
Renta media	18.138	18.309	17.743
Índice de Posicionamiento de la Media (%)	100,00%	100,95%	97,83%
Renta Total de la clase (Millones de euros)	840.767,5	591.097,2	249.670,4
Participación de la clase en la Renta Total (%)	100,00%	70,30%	29,70%
Índice de Gini	0,3233	0,3179	0,3354
Tasa de Pobreza (umbral de pobreza: 9.537€)	0,1980	0,1839	0,2304
Brecha Relativa de Pobreza UE	0,3012	0,2952	0,3095

	España en su conjunto	Hombre	Mujer
Resultados 2011 base 2013 Corregidos			
Número de Hogares representados en la muestra	18.186.339	11.851.969	6.334.370
% de hogares representados	100,00%	65,17%	34,83%
Renta media	20.576	20.815	20.026
Índice de Posicionamiento de la Media (%)	100,00%	101,16%	97,33%
Renta Total de la clase (Millones de euros)	961.623,0	677.599,8	284.023,2
Participación de la clase en la Renta Total (%)	100,00%	70,46%	29,54%
Índice de Gini	0,3525	0,3504	0,3571
Tasa de Pobreza (umbral de pobreza: 10.415€)	0,1990	0,1874	0,2254
Brecha Relativa de Pobreza UE	0,2729	0,2492	0,3118

Fuente: Elaboración propia a partir de los datos de las ECV.

La corrección de las ECV base 2008 con respecto de las ECV base 2004, sólo modifica los niveles de renta de estas categorías, pero no modifica prácticamente nada ni la posición relativa de las categorías con respecto de la media nacional, ni los niveles de desigualdad internos en éstas, ni sus tasas de pobreza; pero sí reducen las correspondientes Brechas de pobreza relativa, aunque conservándose el orden de las mismas.

Al aplicar sobre los datos de las ECV base 2013 la metodología de corrección aquí propuesta para compatibilizar los agregados con la CNE, se observa lo que creemos es una importante anomalía. Nos referimos a la que creemos excepcional estimación de renta media en 2008 para el grupo de hogares con sustentadores principales *mujeres*, para el que el procedimiento asigna una renta media de 20.740€, algo mayor que a los hogares con sustentador principal *hombre* (20.573€). De este modo, creemos que esta estimación muestra otra vez una debilidad del procedimiento, que ya hemos comentado anteriormente: la escala de medida en la que se manifiestan los estimadores de incongruencia está ponderando excesivamente los posibles atípicos y, muy especialmente, cuando los colectivos son poco numerosos. Pero, como también hemos ya comentado anteriormente en alguna otra ocasión, este aspecto requiere una mayor investigación futura más profunda.

Con respecto del resto de indicadores, a pesar del aumento general de rentas que la corrección induce en las clases, se respeta fundamentalmente el espíritu de las cifras de los datos de partida: mismos órdenes de los índices de Gini, aunque como generalmente viene ocurriendo en todos los cuadros, un tanto mayores; mismos órdenes de las tasas y Brechas relativas de pobreza; muy parecidas participaciones relativas de las clases en la Renta Total, y de los niveles y posiciones relativas de las rentas medias, con la importante excepción citada.

A continuación, en el cuadro 4.27.a y el cuadro 4.27.b podemos ver el efecto comparativo de las correcciones realizadas en 2008 y 2011, respectivamente, sobre la Renta Disponible Neta per cápita equivalente de los hogares, según el nivel de estudios del sustentador principal del hogar.

Cuadro 4.27.a
**RENTA DISPONIBLE NETA PER CÁPITA EQUIVALENTE,
 SEGÚN NIVEL DE ESTUDIOS DEL SUSTENTADOR PRINCIPAL DEL HOGAR (2008)**

	España en su conjunto	Sin estudios	Primaria	Secundaria, 1ª etapa	Secundaria, 2ª etapa	Estudios Superiores y asimilados
Resultados 2008 base 2004						
Número de Hogares representados en la muestra	17.279.282	436.728	3.106.855	3.991.577	3.059.186	2.567.720
% de hogares representados	100,00%	2,53%	17,98%	23,10%	17,70%	14,86%
Renta media	16.415	11.692	16.807	15.910	17.278	18.324
Índice de Posicionamiento de la Media (%)	100,00%	71,23%	102,39%	96,93%	105,26%	111,63%
Renta Total de la clase (Millones de euros)	752.938,6	13.236,3	134.869,3	197.125,9	169.063,1	121.900,7
Participación de la clase en la Renta Total (%)	100,00%	1,76%	17,91%	26,18%	22,45%	16,19%
Índice de Gini	0,3053	0,3208	0,3133	0,3398	0,2874	0,2920
Tasa de Pobreza (umbral de pobreza: 8.845€)	0,1800	0,3622	0,2011	0,2228	0,1544	0,1194
Brecha Relativa de Pobreza UE	0,2927	0,4126	0,2947	0,3315	0,3026	0,3043

	España en su conjunto	Sin estudios	Primaria	Secundaria, 1ª etapa	Secundaria, 2ª etapa	Estudios Superiores y asimilados
Resultados 2008 base 2013						
Número de Hogares representados en la muestra	17.279.282	404.206	3.046.454	4.017.637	3.117.960	2.601.672
% de hogares representados	100,00%	2,34%	17,63%	23,25%	18,04%	15,06%
Renta media	18.982	12.535	17.711	18.277	20.718	22.021
Índice de Posicionamiento de la Media (%)	100,00%	66,03%	93,30%	96,29%	109,15%	116,01%
Renta Total de la clase (Millones de euros)	870.712,0	12.024,1	138.604,4	227.422,8	208.171,7	149.445,7
Participación de la clase en la Renta Total (%)	100,00%	1,38%	15,92%	26,12%	23,91%	17,16%
Índice de Gini	0,3102	0,3145	0,3098	0,3396	0,2938	0,3001
Tasa de Pobreza (umbral de pobreza: 10.024€)	0,1846	0,3874	0,2277	0,2351	0,1422	0,1109
Brecha Relativa de Pobreza UE	0,2581	0,3528	0,2992	0,2893	0,2556	0,2683

	España en su conjunto	Sin estudios	Primaria	Secundaria, 1ª etapa	Secundaria, 2ª etapa	Estudios Superiores y asimilados
Resultados 2008 base 2013 Corregidos						
Número de Hogares representados en la muestra	17.324.391	405.031	3.052.963	4.027.504	3.126.019	2.608.809
% de hogares representados	100,00%	2,34%	17,62%	23,25%	18,04%	15,06%
Renta media	20.618	15.483	20.196	20.715	22.845	23.777
Índice de Posicionamiento de la Media (%)	100,00%	75,09%	97,95%	100,47%	110,80%	115,32%
Renta Total de la clase (Millones de euros)	948.095,4	14.876,0	158.370,3	258.367,0	230.110,1	161.805,5
Participación de la clase en la Renta Total (%)	100,00%	1,57%	16,70%	27,25%	24,27%	17,07%
Índice de Gini	0,3551	0,4174	0,3460	0,3559	0,3368	0,3415
Tasa de Pobreza (umbral de pobreza: 10.535€)	0,1890	0,3460	0,1935	0,1825	0,1401	0,1296
Brecha Relativa de Pobreza UE	0,2698	0,3922	0,2876	0,2691	0,2533	0,2620

Fuente: Elaboración propia a partir de los datos de las ECV.

Cuadro 4.27.b
**RENTA DISPONIBLE NETA PER CÁPITA EQUIVALENTE,
 SEGÚN NIVEL DE ESTUDIOS DEL SUSTENTADOR PRINCIPAL DEL HOGAR (2011)**

	España en su conjunto	Sin estudios	Primaria	Secundaria, 1ª etapa	Secundaria, 2ª etapa	Estudios Superiores y asimilados
Resultados 2011 base 2004						
Número de Hogares representados en la muestra	18.034.410	275.743	2.815.607	4.169.608	3.444.739	2.778.943
% de hogares representados	100,00%	1,53%	15,61%	23,12%	19,10%	15,41%
Renta media	15.867	9.798	15.560	15.263	16.168	18.017
Índice de Posicionamiento de la Media (%)	100,00%	61,75%	98,07%	96,20%	101,90%	113,55%
Renta Total de la clase (Millones de euros)	735.508,7	7.015,2	107.049,1	191.914,7	170.443,8	124.332,8
Participación de la clase en la Renta Total (%)	100,00%	0,95%	14,55%	26,09%	23,17%	16,90%
Índice de Gini	0,3231	0,3759	0,3455	0,3440	0,3289	0,3139
Tasa de Pobreza (umbral de pobreza: 8.445€)	0,2009	0,4508	0,2446	0,2473	0,2127	0,1445
Brecha Relativa de Pobreza UE	0,3224	0,4760	0,3324	0,3515	0,3179	0,2945

	España en su conjunto	Sin estudios	Primaria	Secundaria, 1ª etapa	Secundaria, 2ª etapa	Estudios Superiores y asimilados
Resultados 2011 base 2013						
Número de Hogares representados en la muestra	18.034.410	270.642	2.848.508	4.144.586	3.484.729	2.795.486
% de hogares representados	100,00%	1,50%	15,79%	22,98%	19,32%	15,50%
Renta media	18.138	11.421	16.620	17.729	18.591	21.257
Índice de Posicionamiento de la Media (%)	100,00%	62,97%	91,64%	97,75%	102,50%	117,20%
Renta Total de la clase (Millones de euros)	840.767,5	7.718,4	116.181,3	221.138,4	199.674,4	147.689,4
Participación de la clase en la Renta Total (%)	100,00%	0,92%	13,82%	26,30%	23,75%	17,57%
Índice de Gini	0,3233	0,3524	0,3209	0,3496	0,3223	0,3311
Tasa de Pobreza (umbral de pobreza: 9.537€)	0,1980	0,4422	0,2328	0,2512	0,1963	0,1503
Brecha Relativa de Pobreza UE	0,3012	0,3987	0,3690	0,3091	0,2702	0,2811

	España en su conjunto	Sin estudios	Primaria	Secundaria, 1ª etapa	Secundaria, 2ª etapa	Estudios Superiores y asimilados
Resultados 2011 base 2013 Corregidos						
Número de Hogares representados en la muestra	18.186.339	272.637	2.871.401	4.177.132	3.514.311	2.819.870
% de hogares representados	100,00%	1,50%	15,79%	22,97%	19,32%	15,51%
Renta media	20.576	13.595	19.286	20.515	22.494	23.538
Índice de Posicionamiento de la Media (%)	100,00%	66,07%	93,74%	99,70%	109,33%	114,40%
Renta Total de la clase (Millones de euros)	961.623,0	9.249,7	135.880,3	257.890,1	243.605,0	164.919,3
Participación de la clase en la Renta Total (%)	100,00%	0,96%	14,13%	26,82%	25,33%	17,15%
Índice de Gini	0,3525	0,3407	0,3284	0,3617	0,3689	0,3602
Tasa de Pobreza (umbral de pobreza: 10.415€)	0,1990	0,3815	0,2069	0,2254	0,1807	0,1557
Brecha Relativa de Pobreza UE	0,2729	0,4073	0,3093	0,2975	0,2786	0,2820

Fuente: Elaboración propia a partir de los datos de las ECV.

Observamos que, la base 2013 penalizó en sus correcciones especialmente a la categoría de hogares cuyo sustentador principal tenía nivel de estudios de *primaria*, haciendo descender su posición relativa en torno al 91%-93% de la media nacional, según el año. A este respecto, la aplicación de nuestra metodología incrementa lógicamente las rentas para buscar la congruencia con la CNE así como sus posiciones relativas con respecto de la media nacional, excepto para los hogares cuyo sustentador principal tienen nivel de estudios *superiores* o *asimilados*, que bajan un poco estas posiciones hasta el 115,32% en 2008, y hasta el 114,40% en 2011.

Con respecto de los indicadores de desigualdad y las tasas de pobreza, la base 2013 mantiene la tónica de la base 2004. La nueva metodología incrementa los indicadores de desigualdad, con muy ligeros cambios en el orden de las categorías; y hace bajar las tasas de pobreza manteniendo el orden, con la excepción de la categoría superior, en la que aumenta ligeramente, manteniéndose realmente baja (en torno al 13%-15%). Por su parte, las Brechas relativas de pobreza con el método propuesto permanece en el mismo orden que las de la base 2013, que fueron sustancialmente menores que las dadas por la base 2004.

Por último, en el cuadro 4.28.a y el cuadro 4.28.b podemos ver el efecto comparativo de las correcciones realizadas en 2008 y 2011, respectivamente, sobre la Renta Disponible Neta per cápita equivalente de los hogares, según la nacionalidad del sustentador principal del hogar.

Se observa en ellos que la base 2013 reduce de forma significativa la renta media del resto de países de la UE, presumiblemente, debido a la incorporación en la UE de nuevos países no integrados inicialmente en la base 2004, ya que en 2001 no se aprecia este efecto. Pero en ambos años, la base 2013 baja sustancialmente la posición relativa de los hogares cuyo sustentador principal es de alguno de los países del resto del mundo con respecto de la renta media del conjunto de hogares. A este respecto, la metodología propuesta mantiene la tónica de la base 2013, pero recuperando algo de renta en esta misma categoría de hogares.

Los índices de Gini, muy corregidos al alza por la base 2013 sólo para la categoría de Resto de la UE, se ven nuevamente incrementados por el método propuesto pero manteniendo el orden en todo caso.

Cuadro 4.28.a
**RENTA DISPONIBLE NETA PER CÁPITA EQUIVALENTE,
 SEGÚN NACIONALIDAD DEL SUSTENTADOR PRINCIPAL DEL HOGAR (2008)**

	España en su conjunto	Española	Resto de la UE	Resto del mundo
Resultados 2008 base 2004				
Número de Hogares representados en la muestra	17.279.282	15.340.149	336.189	1.314.294
% de hogares representados	100,00%	88,78%	1,95%	7,61%
Renta media	16.415	17.153	16.628	10.129
Índice de Posicionamiento de la Media (%)	100,00%	104,50%	101,30%	61,70%
Renta Total de la clase (Millones de euros)	752.938,6	680.698,4	13.437,1	43.562,2
Participación de la clase en la Renta Total (%)	100,00%	90,41%	1,78%	5,79%
Índice de Gini	0,3053	0,2958	0,3000	0,3351
Tasa de Pobreza (umbral de pobreza: 8.845€)	0,1800	0,1505	0,1396	0,4692
Brecha Relativa de Pobreza UE	0,2927	0,2702	0,3007	0,3469

	España en su conjunto	Española	Resto de la UE	Resto del mundo
Resultados 2008 base 2013				
Número de Hogares representados en la muestra	17.279.282	15.362.867	584.942	1.055.249
% de hogares representados	100,00%	88,91%	3,39%	6,11%
Renta media	18.982	20.038	13.438	10.097
Índice de Posicionamiento de la Media (%)	100,00%	105,56%	70,79%	53,19%
Renta Total de la clase (Millones de euros)	870.712,0	796.094,6	22.843,2	34.586,1
Participación de la clase en la Renta Total (%)	100,00%	91,43%	2,62%	3,97%
Índice de Gini	0,3102	0,2962	0,3766	0,3070
Tasa de Pobreza (umbral de pobreza: 10.024€)	0,1846	0,1433	0,4004	0,5660
Brecha Relativa de Pobreza UE	0,2581	0,2189	0,3673	0,3516

	España en su conjunto	Española	Resto de la UE	Resto del mundo
Resultados 2008 base 2013 Corregidos				
Número de Hogares representados en la muestra	17.324.391	15.404.784	586.441	1.056.952
% de hogares representados	100,00%	88,92%	3,39%	6,10%
Renta media	20.618	21.630	14.208	13.133
Índice de Posicionamiento de la Media (%)	100,00%	104,91%	68,91%	63,70%
Renta Total de la clase (Millones de euros)	948.095,4	861.535,3	24.213,2	45.064,5
Participación de la clase en la Renta Total (%)	100,00%	90,87%	2,55%	4,75%
Índice de Gini	0,3551	0,3454	0,4353	0,3718
Tasa de Pobreza (umbral de pobreza: 10.535€)	0,1890	0,1635	0,3581	0,4209
Brecha Relativa de Pobreza UE	0,2698	0,2576	0,3865	0,2891

Fuente: Elaboración propia a partir de los datos de las ECV.

Cuadro 4.28.b
**RENTA DISPONIBLE NETA PER CÁPITA EQUIVALENTE,
 SEGÚN NACIONALIDAD DEL SUSTENTADOR PRINCIPAL DEL HOGAR (2011)**

	España en su conjunto	Española	Resto de la UE	Resto del mundo
Resultados 2011 base 2004				
Número de Hogares representados en la muestra	18.034.410	16.181.845	629.604	1.106.491
% de hogares representados	100,00%	89,73%	3,49%	6,14%
Renta media	15.867	16.612	13.103	8.584
Índice de Posicionamiento de la Media (%)	100,00%	104,70%	82,58%	54,10%
Renta Total de la clase (Millones de euros)	735.508,7	677.456,3	21.545,6	30.108,2
Participación de la clase en la Renta Total (%)	100,00%	92,11%	2,93%	4,09%
Índice de Gini	0,3231	0,3081	0,3661	0,3698
Tasa de Pobreza (umbral de pobreza: 8.445€)	0,2009	0,1636	0,3433	0,5724
Brecha Relativa de Pobreza UE	0,3224	0,3010	0,3179	0,3947

	España en su conjunto	Española	Resto de la UE	Resto del mundo
Resultados 2011 base 2013				
Número de Hogares representados en la muestra	18.034.410	16.174.778	625.815	1.113.766
% de hogares representados	100,00%	89,69%	3,47%	6,18%
Renta media	18.138	19.074	15.137	8.804
Índice de Posicionamiento de la Media (%)	100,00%	105,16%	83,46%	48,54%
Renta Total de la clase (Millones de euros)	840.767,5	777.119,6	24.748,5	31.215,6
Participación de la clase en la Renta Total (%)	100,00%	92,43%	2,94%	3,71%
Índice de Gini	0,3233	0,3022	0,4603	0,3485
Tasa de Pobreza (umbral de pobreza: 9.537€)	0,1980	0,1542	0,3895	0,6173
Brecha Relativa de Pobreza UE	0,3012	0,2514	0,3960	0,4176

	España en su conjunto	Española	Resto de la UE	Resto del mundo
Resultados 2011 base 2013 Corregidos				
Número de Hogares representados en la muestra	18.186.339	16.311.797	631.157	1.122.283
% de hogares representados	100,00%	89,69%	3,47%	6,17%
Renta media	20.576	21.486	16.805	11.116
Índice de Posicionamiento de la Media (%)	100,00%	104,42%	81,67%	54,03%
Renta Total de la clase (Millones de euros)	961.623,0	882.599,0	27.726,2	39.716,7
Participación de la clase en la Renta Total (%)	100,00%	91,78%	2,88%	4,13%
Índice de Gini	0,3525	0,3361	0,4384	0,3569
Tasa de Pobreza (umbral de pobreza: 10.415€)	0,1990	0,1644	0,3781	0,5117
Brecha Relativa de Pobreza UE	0,2729	0,2556	0,2757	0,3318

Fuente: Elaboración propia a partir de los datos de las ECV.

Se observa que las tasas de pobreza y las Brechas relativas de Pobreza, muy corregidas al alza por la base 2013 en las dos categorías de extranjeros, suavizan dichas correcciones al aplicarles el método propuesto, manteniendo el orden relativo entre las categorías (salvo en el caso del resto de países de la UE en 2008, que ve incrementarse algo su Brecha relativa de pobreza).

A la vista de los resultados expuestos en este apartado, podemos concluir que el procedimiento parece prometedor, respetando bastante el espíritu de las correcciones que realiza la base 2013 mientras estima la distribución de rentas congruente con las cifras de referencia de la CNE, aunque parece que se ha vuelto a presentar una vez más el efecto indeseable que sobre las correcciones parecen tener los datos muy atípicos, cuyo tratamiento en el procedimiento propuesto requieren una investigación más profunda sobre la naturaleza y forma de la ocultación que permitan identificarlos y tratarlos adecuadamente, así como sobre la determinación de una mejor forma funcional (no necesariamente lineal) que permita correcciones oportunas, menos críticas, en presencia de tales atípicos.

Conclusiones y Líneas Futuras de Investigación

Resumidamente, por un lado, se ha puesto de manifiesto que la declaración de ingresos por parte de los hogares cuando se les solicita información en la Encuesta de Condiciones de Vida (ECV) proporciona una distribución personal de la renta sesgada a la baja que difiere de forma significativa de la que verdaderamente se dibuja con los datos oficiales que publica la Contabilidad Nacional de España (CNE), y se han analizado pormenorizadamente estas diferencias mediante la construcción de una cuenta de rentas altamente comparable para ambas fuentes. Por otra parte, dado que existen otras fuentes estadísticas de índole económica dirigidas a hogares, como la Encuesta de Presupuestos Familiares (EPF) y la Encuesta Financiera de las Familias (EFF), se ha procedido a su aprovechamiento mediante la utilización de técnicas de fusión estadística con el objetivo de ampliar con éstas la información de los hogares de la ECV. Así, esta gran base de datos fusionada nos ha permitido la construcción de un conjunto de indicadores que nos informan del nivel de incongruencia en las declaraciones de cada hogar en cada tipo de renta, bien basados en el cumplimiento de las leyes que se deducen de la Teoría Económica, o bien basados en diversas pruebas circunstanciales que se deducen de las propias sinergias derivadas del propio proceso de fusión, que nos dan cierta información de la propensión a la ocultación o sobredeclaración incurrida por cada hogar en las componentes de su renta. Y, finalmente, con base en estos indicadores, hemos propuesto un modelo para la corrección de los diversos componentes de la *Renta disponible de cada uno de los hogares* de la ECV y procedido a su aplicación para la obtención de una distribución personal de la renta, estimación combinada de la distribución personal de la renta, más próxima a la realidad que dibuja la CNE, que la que se deriva directamente de las ECV originales.

A continuación, se desglosan las principales conclusiones obtenidas a lo largo de esta Memoria, organizadas por capítulos, así como las principales *líneas futuras de investigación* que sugieren y quedan abiertas.

En el capítulo I se ha llevado a cabo el estudio de las principales fuentes de información sobre ingresos de los hogares disponibles en España, tanto a nivel individual como agregado, con la finalidad de seleccionar, de entre aquéllas, las más convenientes para el estudio de la distribución personal de la renta que abordamos en este trabajo. Tras un análisis exhaustivo de cada una de ellas, hemos convenido, en primer lugar, que a nivel de microdatos, la fuente que proporciona un mayor grado de detalle sobre la renta de forma homogénea a lo largo del tiempo es la ECV, que se distingue además por pertenecer al conjunto de operaciones estadísticas armonizadas para los países de la Unión Europea y poseer su correspondiente respaldo legal en forma de reglamento comunitario; lo que le confiere una gran claridad a las definiciones y transparencia a los métodos que emplea. Como resultado de este estudio, también nos hemos hecho eco de las posibilidades y oportunidades de utilización conjunta con las ECV que nos proporcionan el resto de fuentes microeconómicas analizadas en el contexto del desarrollo de esta investigación. Por otra parte, a nivel agregado, hemos concluido tomar como referencia los datos oficiales publicados por la CNE, cuyo marco contable, en la actualidad SEC 2010, también es de obligado cumplimiento por los países europeos, propiciándole análogamente claridad y transparencia a sus definiciones y métodos. Por último, una vez analizadas todas las fuentes

de información de interés en este trabajo y el periodo para el que se encuentran disponibles, se han seleccionado para este estudio los años 2008 y 2011 debido a la coexistencia y disponibilidad en estos dos años de todas las encuestas que finalmente nos han sido de utilidad para obtener la estimación combinada de las distribuciones personales de renta en los hogares (ECV, EPF, EFF y CNE) y valorar el comportamiento del método de estimación en diferentes momentos (esos dos años), así como para analizar la robustez de las estimaciones que proporciona cuando se parte de diferentes conjuntos de datos con distinto grado de infradeclaración (la base 2004 y la base 2013 de las ECV).

En lo que concierne a las líneas de investigación futuras sugeridas por lo tratado en este capítulo, hay que decir que el modelo que hemos desarrollado en este trabajo ha sido ejecutado para el Conjunto Nacional como clase única. Sin embargo, su formulación permite, análogamente, ser aplicado para cualquier partición de clases de las que se disponga información agregada de referencia. Y CNE aporta esta información para las regiones españolas a través de su pormenorización regional (CRE). Lamentablemente, en el periodo en que se ha llevado a cabo el desarrollo este trabajo, no estaban disponibles estos datos regionales para el año 2008 en la base 2010 de CNE. Es por ello por lo que en el contexto de este trabajo no lo hemos abordado, ya que era importante considerar los años 2008 y 2011 por los motivos antes expuestos. Una vez que se disponga de la nueva ola de la EFF que tiene como referencia el año 2014 y disponiendo de datos de CRE base 2010 también para dicho año, se podría estudiar la aplicación del procedimiento a nivel regional para los años 2011 y 2014, y analizar análogamente su comportamiento y resultados.

En el capítulo II, se ha constatado, en primer lugar, que si bien las ECV de 2009 y 2012 permiten calcular la *renta disponible de los hogares* para 2008 y 2011, respectivamente, a nivel individual, se aprecian ciertas desviaciones a nivel agregado con respecto de las principales cifras de la CNE, encontrándose comportamientos singularmente llamativos en algunas categorías, especialmente en las de todo tipo de empresarios y profesionales libres. Precisamente, con el objetivo de seguir profundizando en estas diferencias entre las citadas fuentes, se ha procedido al estudio del marco contable teórico de referencia –SEC 2010–, lo que nos ha permitido diseñar una Cuenta de Renta del Sector Hogares (CRSH) para la ECV homogénea a la que conduce a la obtención de la *Renta disponible de los hogares* en CNE. La CRSH así construida para la ECV nos ha permitido detectar que existen partidas no disponibles en la información proporcionada por esta encuesta cuya estimación deberá afrontarse en el desarrollo de este trabajo para poder realizar una comparación lo más homogénea posible de los datos de la ECV con los de la CNE. Una primera estimación de esta CRSH para dicha encuesta, nos permite cuantificar la infradeclaración presente en las diferentes partidas de renta para los años de estudio, muy considerables en las ECV base 2004 en las *Rentas de la propiedad netas recibidas* y en el bloque de *Otras transferencias corrientes tanto percibidas como pagadas*. Además, se ha podido comprobar cómo la incorporación de los registros administrativos en la producción de datos relativos a los ingresos del hogar para las ECV base 2013, ha supuesto no sólo una mayor eficiencia en el método de recogida, sino también una mejora de la calidad de los datos, constatándose una reducción

de los niveles de infradeclaración con respecto a la CNE en todas las partidas que componen la *Renta disponible del hogar*, a excepción de las *Otras transferencias corrientes percibidas* y *pagadas*. Finalmente, los diferentes grados de cobertura detectados en este capítulo entre las partidas que componen la *Renta disponible de los hogares*, manifiestan paralelamente muy distintos grados de ocultación de rentas en las mismas, hecho que podría estar sesgando a la baja las estimaciones de las rentas en aquellos colectivos que se caracterizan por alimentarse principalmente de los tipos de renta en los que más se oculta. Esta asimetría observada en la distribución del desajuste es una prueba adicional de la necesidad de estimar una distribución personal de la renta corregida que tenga en cuenta dicha asimetría.

Este segundo capítulo sugiere algunas líneas de investigación abiertas con relación a la utilización de las verdaderas rentas declaradas por los hogares en las ECV, en lugar de las empleadas en el trabajo que han sido las registradas en los correspondientes archivos de microdatos, que a veces no son exactamente cifras declaradas sino resultantes ciertos procesos de imputación ante la falta de respuesta. Para ello, las variables de tipo “flags” (que se denotan en la encuesta con el sufijo “_F”) y que también suministra las ECV, dan información de la cantidad imputada por el INE, en cada variable y hogar, ante la falta total o parcial de respuesta. Mediante ellas podría resultar muy interesante obtener los datos realmente declarados por los individuos y hogares en los momentos de las encuestas (base 2004) y compararlos con los corregidos mediante su confrontación con los registros administrativos (base 2013). Así se podría disponer de una buena base de datos con la que estudiar más profundamente la caracterización del fenómeno de la ocultación en las declaraciones de rentas y su relación con la tipología de los declarantes.

En el capítulo III, con la finalidad de estimar una distribución personal de la renta más próxima a la realidad que nos muestra la CNE se han estudiado las relaciones que la Teoría Económica nos proporciona entre la *renta disponible de los hogares* y el resto de las principales variables económicas relacionadas con ella, encontrándose entre ellas las relacionadas con el consumo y la riqueza neta del hogar. Apoyándonos en estas relaciones se ha llevado a cabo mediante procesos de “*matching*” el emparejamiento de los hogares de la EPF con los de la ECV y, posteriormente, sobre la base de datos resultante, los hogares de ésta se han fusionado con la información de la EFF, disponiendo de este modo de una estimación sobre el consumo y la situación financiera de las familias de la ECV. Esta gran cantidad de información no ha permitido, por un lado, proceder a la estimación de las partidas de la CRSH, definida en el capítulo anterior, no disponibles en la ECV, teniéndose que alguna de ellas no ha podido ser estimada debido a que no se ha encontrado información para ello en ninguna de las fuentes fusionadas. Es el caso de las indemnizaciones de seguros no vida recibidas por los hogares en compensación de los daños sufridos en sus bienes, que en CNE forman parte de la partida *Indemnizaciones de seguro no vida*, en el bloque de *Otras transferencias corrientes recibidas*. Por esta misma razón, se ha encontrado que las *rentas percibidas por los menores de 16 años* que recoge la ECV tampoco se han podido desagregar en los diferentes componentes de renta en que se han distribuido en dicha Cuenta de Rentas, ya que todas las fuentes recogen la información necesaria sólo para los adultos. Por último, como subproductos de estas fusiones, ha sido definido en este capítulo un conjunto de indicadores de incongruencia para

las declaraciones de los hogares en cada uno de los componentes de su *renta disponible*. La definición de estos se ha realizado bien con base en las relaciones teóricas estudiadas al comienzo de este capítulo, o bien con base en una serie de pruebas circunstanciales derivadas del aprovechamiento de la información fusionada de la EPF y, muy especialmente, de la EFF. Y sus valores pueden emplearse para cuantificar el grado o propensión a la ocultación que tiene de cada hogar en cada componente de su renta. Será en el próximo capítulo donde los hemos calculado y empleado para proceder a la corrección de dichas partidas de renta en la ECV.

Como líneas de investigación abiertas en este *Capítulo III*, el hecho de poder disponer de una base de datos del calibre de la obtenida, es decir, con información sobre los ingresos, el consumo y la situación financiera de las familias, nos abre a futuro una vía muy amplia de estudio, no sólo para corregir los componentes de la *renta disponible de los hogares* infraestimados en las ECV -tal y como hemos llevado a cabo en esta investigación-, sino también, para conocer la utilización de la renta por dichas unidades económicas, es decir, la parte de ésta destinada al *consumo final de bienes y servicios*. A su vez, la parte de la *renta disponible* no consumida por los hogares nos informaría sobre el *ahorro* de éstos, y así podríamos enlazar este estudio con las cuentas de acumulación, cuyo saldo nos proporcionaría la *capacidad/necesidad de financiación* de dicho sector. Para obtener dicho saldo tendríamos información en nuestra base de datos sobre las *transferencias de capital* recibidas y pagadas por los hogares, y sobre la *inversión* que éstos realizan. De este modo, podríamos conocer en qué medida el ahorro de un hogar determinado habría sido suficiente para hacer frente a sus inversiones a lo largo del periodo de estudio. Todo ello, que inicialmente puede intentarse más limitadamente de forma directa con los microdatos de las EFF a nivel nacional, ahora podría abordarse más ampliamente y con mayor riqueza informativa y más posibilidades de segmentación (por ejemplo a nivel regional).

En el *capítulo IV*, con base en los métodos propuestos en (Pena *et al.*, 1996) y (Callealta y Romero, 2013a), se ha desarrollado un nuevo procedimiento para obtener finalmente la pretendida estimación combinada de la distribución personal de rentas en los hogares, basado en la corrección de los microdatos de renta procedentes de las ECV, que hace pivotar las correcciones realizadas, no sobre las declaraciones de los hogares como hacían los métodos anteriores, sino sobre los grados de incongruencia medidos para las mismas. Además, la comparativa de las dos bases de las ECV (2004 y 2013), que ha sido posible mediante técnicas de fusión estadística, ha permitido realizar unas estimaciones objetivas de las progresividades de las ocultaciones que los métodos anteriores dejaban a la elección (parametrización) subjetiva de los investigadores. A este respecto, esta comparación ha permitido también apreciar que el uso de registros administrativos no sólo ha aflorado parte de las esperadas ocultaciones, sino que también ha desvelado la existencia de importantes sobredeclaraciones difícilmente imputables a los efectos de redondeo en las encuestas. Presumiblemente podrían derivarse de un exceso de imputación en aquéllos registros que no daban respuestas o las daban incompletas; lo que debe ser investigado en mayor profundidad. Análogamente a aquellos estudios, las estimaciones obtenidas respetan totalmente los agregados de CNE en todas

las partidas de los ingresos consideradas; pero, comparativamente, las correcciones realizadas sobre los microdatos parecen haberse moderado sustancialmente, no alterando tanto como aquéllos la distribución inicial que mostraban directamente las ECV. Además, el método se muestra robusto en sus estimaciones, que se aproximan considerablemente, a tenor de lo que se observa cuando se compara las correcciones obtenidas partiendo de dos situaciones tan diferentes como las de las dos bases publicadas de las ECV. Quizás se observan correcciones excesivamente importantes para los casos más extremos; lo que debe animar al análisis de la influencia de posibles atípicos en este método.

Consecuentemente, entre las líneas de investigación abiertas en este cuarto capítulo cabe destacar, en primer lugar, que los valores obtenidos para la progresividad de las ocultaciones, "k", con la suposición de una forma lineal para ella, sugieren que debe pensarse más profundamente sobre si puede ser acertada esta suposición en un caso general, para todos los tipos de renta. Es decir, podría investigarse en un futuro cual sería la forma funcional más adecuada para la tasa de corrección de cada tipo de renta, como función de los correspondientes indicadores de incongruencia de las declaraciones. Por otra parte, pensamos que el método funcionaría más realísticamente si las correcciones se hicieran pivotar con un número amplio de indicadores de incongruencia para cada variable a corregir. Si bien cada indicador de incongruencia podría considerarse como una relativamente débil "prueba circunstancial" de la potencial ocultación (o sobredeclaración) de un caso, la consideración conjunta de muchas "pruebas circunstanciales" de este tipo (indicadores) podría configurar pruebas relativamente fuertes de "acusación" de ocultación de los casos, cuando todas apuntan a ello. En este sentido, sería conveniente construir baterías más amplias de tales indicadores de incongruencia para las cantidades declaradas de cada tipo.

Bibliografía

- ACKLEY, G.** (1965): *Teoría Macroeconómica*. México, Unión Tipográfica Editorial Hispano Americana (U.T.E.H.A.).
- ALCAIDE INCHAUSTI, J. Y ALCAIDE GUINDO, P.** (2000): *Renta nacional de España y su distribución provincial. Año 1995 y avances 1996-1999*. Bilbao, Fundación BBVA.
- ALCAIDE INCHAUSTI, A. Y ALCAIDE INCHAUSTI, J.** (1974): "Metodología para la estimación de la distribución personal de la renta en España en 1970". *Hacienda Pública Española*, 26, pp. 55-63.
- (1977): "Distribución personal de la renta en España y en los países de la OCDE". *Hacienda Pública Española*, 47, pp. 17-57.
- (1984): "Distribución personal de la renta española en 1980". *Hacienda Pública Española*, 85, pp. 485-509.
- ÁLVAREZ, C., AYALA, L., IRIONDO, I., MARTÍNEZ, R., PALACIO, J. Y RUIZ-HUERTA, J.** (1996): *La distribución funcional y personal de la renta en España. Un análisis de sus relaciones*. Consejo Económico y Social (CES), Colección de Estudios, 30, Madrid.
- ANDO, A. AND MODIGLIANI, F.** (1963): "The 'life-cycle' hypothesis of saving: Aggregate implications and tests". *American Economic Review*, Vol. 53, No. 1, pp. 55-84.
- ANDRÉS, L. Y MERCADER, M.** (2001): "Sobre la fiabilidad de los datos de renta en el Panel de Hogares de la Unión Europea (PHOGUE, 1994)". *Estadística Española*, Vol. 43, No. 148, pp. 241-280.
- ARCE, Ó., PRADES, E., Y URTASUN, A.** (2013): "La evolución del ahorro y del consumo de los hogares españoles durante la crisis". *Boletín Económico*, septiembre 2013, pp. 65-73. Banco de España. Recuperado de: <http://www.bde.es/f/webbde/SES/Secciones/Publicaciones/InformesBoletinesRevistas/BoletinEconomico/13/Sep/Fich/be1309.pdf>
- ARGANDOÑA, A.** (1986): "Los determinantes del ahorro de las familias". *Papeles de Economía Española*, 28, pp. 317-361.
- ARGIMÓN, I., GONZÁLEZ-PÁRAMO, J.M. Y ROLDÁN, J.M.** (1992): "Ahorro, riqueza y tipos de interés en España". Documento de Trabajo No. 9215. Servicio de Estudios. Banco de España. Recuperado de: http://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosTrabajo/92/Fich/dt_9215.pdf
- ATKINSON, A. B.** (1970): "On the Measurement of Inequality". *Journal of Economic Theory*, Vol. 2, No. 3, pp. 244-263. ([http://dx.doi.org/10.1016/0022-0531\(70\)90039-6](http://dx.doi.org/10.1016/0022-0531(70)90039-6))
- (1987): "On the Measurement of Poverty". *Econometrica*, Vol. 55, No. 4, pp. 749-764. (<http://dx.doi.org/10.2307/1911028>)
- AUSTRALIAN BUREAU OF STATISTICS** (2003): "Revised household income distribution statistics", in *Australian economic indicators*, June, Catalogue No. 1350.0, pp. 3-15. Canberra, ABS.

- AYALA, L., MARTÍNEZ, R. Y RUIZ-HUERTA, J.** (1996): "La distribución de la renta en España desde una perspectiva internacional: Tendencias y factores de cambio", en *La desigualdad de recursos (II Simposio sobre desigualdad y distribución de la renta y la riqueza)*, Fundación Argentaria.
- AYALA, L. Y ONRUBIA, J.** (2001): "La Distribución de la Renta en España según Datos Fiscales". *Papeles de Economía Española*, 88, pp. 89-112.
- BAKER, K., HARRIS, P. & O'BRIEN, J.** (1989): "Data fusion: An appraisal and experimental evaluation". *Journal of the market research society*, vol. 31, No. 2, pp.153-212.
- BANCO BILBAO VIZCAYA (BBV)** (1984): *Renta Nacional de España y su distribución provincial 1981*. Bilbao, BBV.
- BANCO DE ESPAÑA** (2004): "Encuesta Financiera de las Familias (EFF): Descripción, métodos y resultados preliminares". *Boletín Económico*, noviembre 2004, pp. 61-82. Banco de España. Recuperado de: <http://www.bde.es/f/webbde/Secciones/Publicaciones/InformesBoletinesRevistas/BoletinEconomico/04/Fich/be0411.pdf>
- (2007): "Encuesta Financiera de las Familias (EFF) 2005: Métodos, resultados y cambios entre 2002 y 2005". *Boletín Económico*, diciembre 2007, pp. 31-66. Banco de España. Recuperado de: <http://www.bde.es/f/webbde/Secciones/Publicaciones/InformesBoletinesRevistas/BoletinEconomico/07/Fich/be0712.pdf>
- (2010a): "Encuesta Financiera de las Familias (EFF) 2008: Métodos, resultados y cambios desde 2005". *Boletín Económico*, diciembre 2010, pp. 29-64. Banco de España. Recuperado de: <http://www.bde.es/f/webbde/SES/Secciones/Publicaciones/InformesBoletinesRevistas/BoletinEconomico/10/Dic/Fich/be1012.pdf>
- (2010b): *Encuesta Financiera de las Familias. Microdatos Año 2008*. [En línea]. Recuperado de: http://www.bde.es/bde/es/areas/estadis/Otras_estadistic/Encuesta_Financi/
- (2014a): "Encuesta Financiera de las Familias (EFF) 2011: Métodos, resultados y cambios desde 2008". *Boletín Económico*, enero 2014, pp. 71-104. Banco de España. Recuperado de: <http://www.bde.es/f/webbde/SES/Secciones/Publicaciones/InformesBoletinesRevistas/BoletinEconomico/14/Ene/Fich/be1401.pdf>
- (2014b): *Encuesta Financiera de las Familias. Microdatos Año 2011*. [En línea]. Recuperado de: http://www.bde.es/bde/es/areas/estadis/Otras_estadistic/Encuesta_Financi/
- (2017): "Encuesta Financiera de las Familias (EFF) 2014: Métodos, resultados y cambios desde 2011". *Artículos analíticos*, enero 2017, Banco de España. Recuperado de: <http://www.bde.es/f/webbde/SES/Secciones/Publicaciones/InformesBoletinesRevistas/ArticulosAnaliticos/2017/T1/fich/beaa1701-art2.pdf>

- (2017): "Tipos de interés aplicados por las instituciones financieras monetarias a residentes en la UEM (CBE 1/2010). Datos desde enero de 2003. Tipos de interés (TAE y TEDR) de nuevas operaciones. Depósitos de hogares e ISFLSH y sociedades no financieras". Recuperado de: <http://www.bde.es/webbde/es/estadis/infoest/series/be1908.csv>
- BARBANCHO, A.G.** (1961): "La función de consumo de la economía española". *Revista de Estudios Agrosociales*, 37. Recuperado de: http://www.mapama.gob.es/ministerio/pags/Biblioteca/Revistas/pdf_reas%2Fr037_03.pdf
- BEATON, A. E.** (1975): "The influence of education and ability on salary and attitudes", in F. Thomas Juster (Ed.): *Education, Income, and Human Behavior*, pp. 365-396. New York, McGraw-Hill. Recuperado de: <http://www.nber.org/chapters/c3704.pdf>
- BECKER, G. S.** (1964): *Human capital. First edition*. New York, Columbia University Press for the National Bureau of Economic Research.
- BLANCO, M. A.; FRUTOS, R.; LÓPEZ, E Y MARTÍNEZ, S.** (2014): *La influencia de la riqueza financiera sobre el consumo privado: Análisis de la experiencia española durante la "Gran Recesión"*. Documento de trabajo 03/2014. Instituto Universitario de Análisis Económico y Social, Universidad de Alcalá. Recuperado de: http://www3.uah.es/iaes/publicaciones/DT_03_14.pdf
- BOHEIM, R. AND JENKINS, S. P.** (2000): *Do current income and annual income measures provide different pictures of Britain's income distribution?* ISER Working Paper Series No. 2000-16. Institute for Social and Economic Research (ISER), University of Essex, Colchester, UK. Recuperado de: <https://www.iser.essex.ac.uk/research/publications/working-papers/iser/2000-16.pdf>
- BORRÁS, V.** (2007): "Las desigualdades en el consumo a través del género". *Revista Española de Sociología*, 8, pp. 139-156. Recuperado de: <http://www.fes-sociologia.com/files/res/8/06.pdf>
- BOSCH, A., ESCRIBANO, C. Y SÁNCHEZ, I.** (1989): *Evolución de la desigualdad y la pobreza en España: Estudio basado en las encuestas de presupuestos familiares 1973-74 y 1980-81*. Madrid, INE.
- BOVER, O.** (2004): *The Spanish Survey of Household Finances (EFF): Description and methods of the 2002 wave*. Documentos de Trabajo, No. 0409, Banco de España. Recuperado de: <http://www.bde.es/f/webbde/SES/Secciones/Publicaciones/Publicaciones/Seriadas/DocumentosOcasionales/04/Fic/do0409e.pdf>
- (2005): *Wealth effects on consumption: Microeconomic estimates from the Spanish Survey of Household Finances*. Documentos de Trabajo, No. 0522, Banco de España. Recuperado de: <http://www.bde.es/f/webbde/SES/Secciones/Publicaciones/Publicaciones/Seriadas/DocumentosTrabajo/05/Fic/dt0522e.pdf>
- BRAYTON, F. AND MAUSKOPF, E.** (1985): "The Federal Reserve Board MPS Quarterly Econometric Model of the U.S. Economy". *Economic Modelling*, Vol. 2, No. 3, pp. 170-292.

- (1987): "Structure and Uses of the MPS Quarterly Econometric Model of the United States". *Federal Reserve Bulletin*, 73, pp. 93-109.
- CAGAN, P.** (1990): "The 1987 Stock Market Crash and the Wealth Effect", en P. A. Klein [ed.]: *Analyzing modern business cycles: Essays honoring Geoffrey Moore*, pp. 249-60. Sharpe, London.
- CAGE, R.** (1989): "Spending differences across occupational fields". *Monthly Labor Review*, Vol. 112, No. 2, pp. 33-43. Recuperado de: <https://www.bls.gov/opub/mlr/1989/12/art5full.pdf>
- CALONGE, S. Y MANRESA, A.** (1997): "Consecuencias redistributivas del Estado del Bienestar en España: Un análisis empírico desagregado". *Moneda y Crédito*, 204, pp. 13-65.
- (2001): "Fusión estadística de bases de datos", en S. Calonge y A. Manresa. (Ed.), *Incidencia fiscal y del gasto público social sobre la distribución de la renta en España y sus CCAA*, pp. 191-246. Bilbao, Fundación BBVA.
- CALLEALTA, F.J. Y GARCÍA, N.** (2011): "Sobre congruencia entre las rentas de la ECV y la CRE". *Anales de Economía Aplicada, ASEPELT España. Reunión anual (2011. Santander)*. pp. 5-21. Recuperado de: <http://www.asepelt.org/ficheros/File/Anales/2011-Santander/ANALES-ECONOMIA-APLICADA-2011.pdf>
- CALLEALTA, F.J. Y ORTEGA, J.** (2002): "Ocultación de los ingresos personales en España: Una aproximación mediante procedimientos robustos". *Estadística Española*, Vol. 44, No. 149, pp. 47-67.
- CALLEALTA, F.J. Y ROMERO, C.** (2013a): "Sobre la congruencia de los datos de renta de las encuestas de condiciones de vida y los de contabilidad nacional". *Revista Galega de Economía*, Vol. 22, No. extraordinario (septiembre 2013), pp. 85-114. Recuperado de: http://www.usc.es/econo/RGE/Vol22_ex/castelan/art4c.pdf
- (2013b): "Una corrección de la distribución de ingresos de los hogares españoles procedente de la E.C.V. mediante técnicas de fusión estadística". *Anales de Economía Aplicada. Actas XXVII Reunión de l'Assoc. Scientifique Européenne pour la Prediction a Long Terme (España)*, pp. 1741-1761. Recuperado de: <http://www.asepelt.org/ficheros/File/Anales/2013/anales-2013.pdf>
- (2016): "La ocultación de ingresos en las E.C.V. y los registros administrativos", en J. Castejón y J.M. Herrerías (Coord.), *Investigaciones en Métodos Cuantitativos para la Economía y la Empresa*, pp. 87-120. Granada, Editorial Universidad de Granada.
- CAMPBELL, J. Y. AND COCCO, J. F.** (2007): "How do house prices affect consumption? Evidence from micro data". *Journal of Monetary Economics*, Vol. 54, No. 3, pp. 591-621. (<http://dx.doi.org/10.1016/j.jmoneco.2005.10.016>)
- CANTÓ, O.** (1997): "Desempleo y pobreza en la España de los 90". *Papeles de Economía Española*, 78, pp. 88-108.

- CANTÓ, O., DEL RÍO, C. Y GRADÍN, C. (2000): "La situación de los estudios sobre desigualdad y pobreza en España". *Cuadernos de Gobierno y Administración*, 2, pp. 25-94. Recuperado de: <http://cgradin.webs.uvigo.es/Publicacions/Estudios%20desigualdad.PDF>
- CAÑADA, A. (1999): "La distribución de la renta: una nota metodológica sobre las nuevas posibilidades de estudio en el SEC95". *Estadística Española*, Vol. 41, No. 144, pp. 203-240.
- CARROL, C. (2000): "Why do the rich save so much?" In J.Slemrod (Ed.), *Does Atlas Shrug? The Economic Consequences of Taxing the Rich*. Cambridge, MA: Harvard University Press. (<http://dx.doi.org/10.3386/w6549>)
- CHARRON-CHÉNIER, R., FINK, J.J. & KEISTER, L. A. (2017): "Race and Consumption: Black and White Disparities in Household Spending". *Sociology of Race and Ethnicity*, Vol. 3, No.1, pp. 50-67.
- CHISWICK, B. R. (1974): *Income inequality: Regional analyses within a human capital framework*. New York, National Bureau of Economic Research.
- CHRISTENSEN, L. (2012): "How Are Interest Rates Affecting Household Consumption and Savings?" *All Graduate Plan B and other Reports*, 215. Recuperado de: <http://digitalcommons.usu.edu/gradreports/215>
- CLINTON, P. M. (2014): "Integration of Micro and Macro Data on Consumer Income and Expenditures", in Dale W. Jorgenson, J. Steven Landefeld, and Paul Schreyer (Editors): *Measuring Economic Sustainability and Progress*, pp. 137-179. Chicago, University of Chicago Press. Recuperado de: <http://www.nber.org/chapters/c12831.pdf>
- COLI, A. AND TARTAMELLA, F. (2012): "Analyzing the Distribution of Income: How to Account for the Underground Economy in the Household Income Micro Datasets". *IARIW 2012 Conference*. Recuperado de: <http://www.iariw.org/papers/2012/TartamellaPaper.pdf>
- CONSOLINI, P. (2009): "Integrazione di dati campionari EU-SILC con dati di fonte amministrativa". *Collana Metodi e Norme Istat*, 38.
- CORTÁZAR, J. (1996): "Regional Effects on the Consumption Decision: Evidence from the Spanish Family Expenditures Survey". Working Paper 9605, CEMFI.
- COWELL, F. (2007): *Income distribution and inequality*. DARP 94. Suntory and Toyota International Centres for Economics and Related Disciplines, London School of Economics and Political Science, London, UK. Recuperado de: [http://eprints.lse.ac.uk/3780/1/Income Distribution and Inequality.pdf](http://eprints.lse.ac.uk/3780/1/Income%20Distribution%20and%20Inequality.pdf)
- CRISTIA, J. AND SCHWABISH, J. (2009): "Measurement error in the SIPP: Evidence from administrative matched records". *Journal of Economic and Social Measurement*, 34, pp. 1-17. Recuperado de: https://www.cbo.gov/sites/default/files/110th-congress-2007-2008/workingpaper/2007-03_0.pdf
- CRISTÓBAL, A. (2010): "El gasto en consumo final en las cuentas nacionales". *Indice: Revista de estadística y sociedad*, 38, pp. 6-8. Recuperado de: <http://www.revistaindice.com/numero38/p6.pdf>

- CRISTOFANI, G. L. (2011): *Los determinantes para el ahorro en los hogares de bajos ingresos de GBA según los enfoques teóricos*. [Tesis de Maestría en Economía Urbana]. Universidad Torcuatto di tella. Recuperado de: http://www.reporteinmobiliario.com/nuke/tesis/guillermo_cristofani_di_tella.pdf
- CUADRADO, J. R., MANCHA, T., VILLENA, J. E., CASARES, J., GONZÁLEZ, M., MARÍN, J. M., Y PEINADO, M. L. (2010): *Política económica: elaboración, objetivos e instrumentos* (4ª Edición). Madrid, McGraw-Hill.
- DAUPHIN, L. (2009): "La qualité des données sur les revenus: Enquête versus fichier administratif". *Les Journées de Méthodologie Statistique*. INSEE.
- DAYAL, N., J. GOMULKA, L. MITTON, H. SUTHERLAND & R. TAYLOR (2000): "Enhancing Family Resources Survey income data with expenditure data from the Family Expenditure Survey: Data comparisons", *University of Cambridge Department of Applied Economics Microsimulation Unit Research Notes*, n.º MU/RN/40. Recuperado de: <https://www.iser.essex.ac.uk/files/msu/publications/pdf/rn40.pdf>
- DEATON, A., RUIZ-CASTILLO, J. Y THOMAS, D. (1989): "The Influence of household Composition on Household Expenditure Patterns: Theory and Spanish Evidence". *Journal of Political Economy*, Vol 97, No. 1, pp. 179-200. (<http://dx.doi.org/10.1086/261597>)
- D'ORAZIO, M., DI ZIO, M., & SCANU, M. (2006): *Statistical matching: Theory and practice*. Chichester, John Wiley & Sons.
- DUESENBERY, J.S. (1949): *Income, saving and the theory of consumer behaviour*. Cambridge, Harvard University Press.
- DUNBAR, G. R. AND FU, C. (2015): *Sheltered income: Estimating income under-reporting in Canada, 1998 and 2004*. Working Paper No. 22. Bank of Canada. Recuperado de: <http://www.bankofcanada.ca/wp-content/uploads/2015/06/wp2015-22.pdf>
- EGUÍA, B. Y ECHEVARRÍA, C. A. (2001): "Estructura de edad poblacional y consumo privado en España"; en FEDEA, *Estudios sobre la Economía Española*, 2001/117. Recuperado de: <http://documentos.fedea.net/pubs/eee/eee117.pdf>
- ELLIOT, J.W. (1980): "Wealth and Wealth Proxies in a Permanent Income Model", *Quarterly Journal of Economics*, 95(3), pp. 509-535. (<http://dx.doi.org/10.2307/1885091>)
- EUROSTAT (1970): *European System of Integrated Economic Accounts (ESA)*: Luxemburgo, Publications Office of the European Union.
- (1979): *European System of Integrated Economic Accounts (ESA second edition)*: Luxemburgo, Publications Office of the European Union.
- (1996): *European System of National and Regional Accounts: ESA 1995*. Luxemburgo, Publications Office of the European Union. Recuperado de: <http://aei.pitt.edu/52185/1/1995.pdf>
- (2001): *Handbook on price and volume measures in national accounts*. Luxemburgo, Publications Office of the European Union. Recuperado de:

- <http://ec.europa.eu/eurostat/ramon/statmanuals/files/KS-41-01-543-EN.pdf>
- (2008): *Eurostat Manual of Supply, Use and Input-Output Tables*. Luxemburgo, Publications Office of the European Union. Recuperado de: <http://ec.europa.eu/eurostat/documents/3859598/5902113/KS-RA-07-013-EN.PDF/b0b3d71e-3930-4442-94be-70b36cea9b39?version=1.0>
 - (2013a): *European System of National and Regional Accounts: ESA 2010*. Luxemburgo, Publications Office of the European Union. Recuperado de: <http://ec.europa.eu/eurostat/documents/3859598/5925693/KS-02-13-269-EN.PDF/44cd9d01-bc64-40e5-bd40-d17df0c69334>
 - (2013b): *The use of registers in the context of EU-SILC: challenges and opportunities*. Luxemburgo, Publications Office of the European Union. Recuperado de: <http://ec.europa.eu/eurostat/documents/3888793/5856365/KS-TC-13-004-EN.PDF/fe05381c-8a77-4aa9-ac23-6c207a3c7f07?version=1.0>
 - (2014a): *Manual on the changes between ESA 95 and ESA 2010*. Luxemburgo, Publications Office of the European Union. Recuperado de: <http://ec.europa.eu/eurostat/documents/3859598/5936825/KS-GK-14-002-EN-N.pdf/059d43eb-2047-40c3-9cdd-d847bfbce33b>
 - (2014b): *The ESS Vision 2020*. Document ESSC 2014/21/7a/EN – 05.05.2014. Recuperado de: <http://ec.europa.eu/eurostat/documents/10186/756730/ESS-Vision-2020.pdf/8d97506b-b802-439e-9ea4-303e905f4255>
- FESSEAU, M. AND MATTONETTI, M. L.** (2013): *Distributional measures across household groups in a National Accounts Framework: Results from an experimental cross-country exercise on household income, consumption and saving*. OECD Statistics Working Papers, No. 2013/04, OECD Publishing. (<http://dx.doi.org/10.1787/5k3wdjqr775f-en>)
- FESSEAU, M., WOLFF, F. & MATTONETTI, M. L.** (2013): *A cross-country comparison of household income, consumption and wealth between Micro Sources and National Accounts Aggregates*. OECD Statistics Working Papers, No. 2013/03, OECD Publishing. (<http://dx.doi.org/10.1787/5k3wdjrn7mv-en>)
- FISHER, I.** (1930): *The Theory of Interest*. New York, Macmillan.
- FRIEND, I. AND LIEBERMAN, CH.** (1975): "Short-Run Asset Effects on Household Saving and Consumption: The Cross-Section Evidence". *American Economic Review*, 65, pp. 624-33.
- FRIEDMAN, M.** (1957): *A theory of consumption function*. Princeton, Princeton University Press.
- GALI, J.** (1990): "Finite horizons, life cycle savings and time series evidence on consumption". *Journal of Monetary Economics*, 26, pp. 433-52. ([http://dx.doi.org/10.1016/0304-3932\(90\)90006-P](http://dx.doi.org/10.1016/0304-3932(90)90006-P))
- GARCÍA, A.** (1980): "Algunos problemas empíricos en las estimaciones de distribuciones de rentas". *Estadística Española*, 89, pp. 57-74.

- GARCÍA, J. Y LABEAGA, J. M.** (1996): "Alternative Approaches to Modelling Zero Expenditure: An Application to Spanish Demand for Tobacco". *Oxford Bulletin of Economics and Statistics*, Vol. 58, No. 3, pp. 489-506. (<http://dx.doi.org/10.1111/j.1468-0084.1996.mp58003004.x>)
- GARDE, J. A., MARTÍNEZ, R. Y RUIZ-HUERTA, J.** (1996): "Los estudios sobre distribución de la renta en España: fuentes, resultados y perspectivas de futuro", en *La desigualdad de recursos. II Simposio sobre Igualdad y Distribución de la Renta y la Riqueza*. Fundación Argentaria, Colección Igualdad, Vol. 6, pp. 257-314.
- GIL, P.** (2003): "Aspectos financieros del consumo de las familias: el efecto riqueza". *Anales de Economía Aplicada, ASEPELT España. Reunión anual* (17. 2003. Almería).
- GOURINCHAS, P. Y PARKER, J. A.** (2002): "Consumption over the life cycle". *Econometrica*, Vol. 70, No.1, pp. 47-89. (<http://dx.doi.org/10.1111/1468-0262.00269>)
- GOWER, J.C.** (1971): "A General Coefficient of Similarity and Some of its Properties". *Biometrics*, 27, pp. 857-871. (<http://dx.doi.org/10.2307/2528823>)
- GRACIA, A., GIL, J.M., Y ANGULO, A. M.** (1998): "El consumo de alimentos en España: el consumidor rural versus urbano". *Revista de estudios regionales*, 50, pp. 111-130. Recuperado de: <https://upcommons.upc.edu/bitstream/handle/2117/2379/6.2.30.pdf>
- GUIO, L., PAIELLA, M. & VISCO, I.** (2005): *Do capital gains affect consumption? Estimates of wealth effects from Italian households' behavior*. Temi di discussione, No. 555, Banca d'Italia. Recuperado de: https://www.bancaditalia.it/pubblicazioni/temi-discussione/2005/2005-0555/tema_555.pdf?language_id=1
- GYLFASON, T.** (1981) "Interest rates, inflation and the aggregate consumption function". *Review of Economics and Statistics*, Vol. 63, No. 2, pp. 233-245. (<http://dx.doi.org/10.2307/1924094>)
- HABERLER, G.** (1939): *Prosperity and Depression, Second edition*. Génova, League of Nations.
- HALL, R.** (1978): "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence". *Journal of Political Economy*, 86, pp. 971-87. (<http://dx.doi.org/10.1086/260724>)
- HAMILTON, D.** (1987): "Institutional economics and consumption". *Journal of Economic Issues*. Vol. 21, No. 4, pp. 1531-1554. (<http://dx.doi.org/10.1080/00213624.1987.11504714>)
- HONKKILA, J. AND I. K. KAVONIUS** (2013): *Micro and macro analysis on household income, wealth and saving in the Euro Area*. Working Papers No. 1619, European Central Bank, Frankfurt am Main 2013. Recuperado de: <https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1619.pdf>
- HOOSAIN, A.** (2012): "Relationship between consumer credit and consumption spending in South Africa". Gordon Institute of Business Science, University of Pretoria. Recuperado de:

<http://repository.up.ac.za/bitstream/handle/2263/30603/dissertation.pdf?sequence=1&isAllowed=y>

HOWARD, M.C. (1979): *Modern Theories of Income Distribution*. New York, McMillan Press.

HURST, E., LI, G., AND PUGSLEY, B. (2014): "Are Household Surveys Like Tax Forms: Evidence from Income Underreporting of the Self-Employed". *Review of Economics and Statistics*. Vol. 96, No. 1, pp. 19-33. (http://dx.doi.org/10.1162/REST_a_00363)

INSTITUTO NACIONAL DE ESTADÍSTICA (INE): "Clasificación de bienes y servicios. COICOP (Classification Of Individual Consumption by Purpose)". Madrid, INE. Recuperado de: <http://www.ine.es/daco/daco42/daco4213/anexoecpf06.pdf>

— (1996): *Panel de Hogares de la Unión Europea (PHOGUE): Metodología*. Madrid, INE.

— (2004): *Clasificaciones utilizadas en la ECV Base 2004*. Madrid, INE. Recuperado de: <http://www.ine.es/daco/daco42/condivi/clasifica.pdf>

— (2005): *Encuesta de Condiciones de Vida: Metodología*. Madrid, INE.

— (2006): *Encuesta de Presupuestos Familiares. Metodología*. Madrid, INE. Recuperado de: <http://www.ine.es/metodologia/t25/t2530p458.pdf>

— (2009): *Encuesta de Presupuestos Familiares. Base 2006. Resultados. Ficheros de Microdatos. Año 2008*. [En línea]. Recuperado de: http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176806&menu=resultados&secc=1254736195147&idp=1254735976608

— (2010): *Encuesta de Condiciones de Vida. Base 2004. Resultados. Microdatos. Ficheros transversales. Año 2009*. [En línea]. Recuperado de: http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176807&menu=resultados&secc=1254736195153&idp=1254735976608

— (2011): *Contabilidad Nacional de España Base 2008. Características metodológicas*. Madrid, INE. Recuperado de: www.ine.es/daco/daco42/cne00/nota_cambmet_b2008.pdf

— (2012a): *Contabilidad Nacional Anual de España. Base 2008. Marco input-output. Tabla de Origen y Destino 2008*. [En línea]. Recuperado de: http://www.ine.es/daco/daco42/cne00/cne_tod_08.xls

— (2012b): *Contabilidad Regional de España. Base 2008. Enfoque institucional. Cuentas de Renta de los Hogares. Serie 2008-2011*. Recuperado de: http://www.ine.es/dynt3/inebase/index.htm?type=pcaxis&path=/t35/p010/_ch_base08&file=pcaxis

— (2012c): *Encuesta de Presupuestos Familiares. Base 2006. Resultados. Ficheros de Microdatos. Año 2011*. [En línea]. Recuperado de: http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176806&menu=resultados&secc=1254736195147&idp=1254735976608

- (2013a): *Clasificaciones utilizadas en la ECV Base 2013*. Madrid, INE. Recuperado de: <http://www.ine.es/daco/daco42/condivi/ecv2013clasifica.pdf>
 - (2013b): *Encuesta de Condiciones de Vida. Base 2004. Resultados. Microdatos. Ficheros transversales. Año 2012*. [En línea]. Recuperado de: http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176807&menu=resultados&secc=1254736195153&idp=1254735976608
 - (2013c): *Encuesta de Condiciones de Vida. Metodología revisada*. Madrid, INE. Recuperado de: http://www.ine.es/daco/daco42/condivi/ecv_metodo.pdf
 - (2014a): *Aprovechamiento de los Ficheros Administrativos en la Encuesta de Condiciones de Vida*. Madrid, INE. Recuperado de: http://www.ine.es/daco/daco42/condivi/ecv_aprov_ficheros.pdf
 - (2014b): *Contabilidad Nacional Anual de España. Base 2008. Cuentas no financieras de los sectores institucionales. Serie 2000-2012*. [En línea]. Recuperado de: http://www.ine.es/daco/daco42/cne08/cecosec00_12.xls
 - (2014c): *Encuesta de Condiciones de Vida. Base 2013. Resultados. Microdatos. Ficheros transversales. Años 2009 y 2012*. [En línea]. Recuperado de: http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176807&menu=resultados&secc=1254736195153&idp=1254735976608
 - (2015): *Cifras de Población a 1 de julio de 2015. Estadística de Migraciones. Primer semestre de 2015. Datos Provisionales*. [En línea]. Recuperado de: http://www.ine.es/daco/inebase_mensual/diciembre_2015/cifras_poblacion.zip
 - (2016a): *Contabilidad Nacional Anual de España. Base 2010. Cuentas no financieras de los sectores institucionales. Serie 1999-2015*. [En línea]. Recuperado de: http://www.ine.es/daco/daco42/cne10/cecosec99_15.xlsx
 - (2016b): *Contabilidad Nacional Anual de España. Base 2010. Resultados detallados. Agregados por ramas de actividad. Serie 1999-2015*. [En línea]. Recuperado de: http://www.ine.es/daco/daco42/cne10/agr_x_ramas95_15.xls
 - (2016c): *Contabilidad Nacional Anual de España. Base 2010. Marco inputoutput. Tabla de Origen y Destino 2011*. [En línea]. Recuperado de: http://www.ine.es/daco/daco42/cne10/cne_tod_11.xlsx.
- INE Y UNIVERSIDAD AUTÓNOMA DE MADRID (1996):** *Encuesta de presupuestos familiares. Desigualdad y pobreza en España: estudio basado en las encuestas de presupuestos familiares de 1973-74, 1980-81 y 1990-91*. Madrid, INE.
- KADANE, J. B. (1978):** "Some Statistical Problems in Merging Data Files". *Compendium of Tax Research*, S. 159-179. Whashington, D.C., U.S. Government Printing Office.

- KAZEMIER, B.** (1990): *Concealed interest income of households in the Netherlands; 1977, 1979 and 1981*. Central Bureau of Statistics. The Netherlands. National Accounts Research Division. Recuperado de: <https://www.cbs.nl/-/media/imported/documents/2006/24/1990-039-p30-pub.pdf?la=nl-nl>
- KEYNES, J. M.** (1936): *The general theory of employment, interest and money*. United States, MacMillan Press.
- KIRAN, T. AND DHAWAN, S.** (2015): "The impact of family size on savings and consumption expenditure of industrial workers: A cross-sectional study". *American Journal of Economics and Business Administration*, Vol. 7, No. 4, pp. 177-184. (<http://dx.doi.org/10.3844/ajebasp.2015.177.184>)
- KLEIN, L. R. AND GOLDBERGER, A. S.** (1955): *An Econometric Model of the United States 1929-1952*. Amsterdam, North Holland.
- KUKK, M. AND STAEHR, K.** (2013): "Income underreporting by households with business income: Evidence from Estonia". *Post-Communist Economies*, Vol. 26, No. 2, pp. 257-276. (<http://dx.doi.org/10.1080/14631377.2014.904110>)
- KUZNETS, S.** (1946): *National Income and its Composition, Vol. II*. New York, NBER.
- LABEAGA, J. M. Y LÓPEZ, A.** (1994): "Estimation of the Welfare Effects of Indirect Tax Changes on Spanish Households: An Analysis of the 1992 VAT Reform". *Investigaciones Económicas*, Vol 18, No. 2, pp. 289-311.
- (1996): "Flexible Demand System Estimation and the Revenue and the Welfare Effects of the 1995 VAT Reform on Spanish Households". *Revista Española de Economía*, Vol. 13, No. 2, pp. 181-197.
- LARRAIN, F. AND SACHS, J.** (2002): *Macroeconomía en la economía global* (2ª. ed.). Buenos Aires, Prentice Hall.
- LENORMAND, T., LOUAIL, T., CANTÚ-ROS, O.G., PICORNELL, M., HERRANZ, R., ARIAS, J.M., BARTHÉLEMY, M., SAN MIGUEL, M. Y RAMASCO, J.J.** (2015): "Influence of sociodemographic characteristics on human mobility". *Scientific Reports* 5, 10075. (<http://dx.doi.org/10.1038/srep10075>)
- LEULESCU, A. AND AGAFITEI, M.** (2012): "A quality framework for matching EU social surveys". Q2012 - *European Conference On Quality In Official Statistics*, Athens 30 May - 1 June 2012. Recuperado de: [http://www.q2012.gr/articlefiles/sessions/24.2 Leulescu-Agafitei_A%20quality%20framework%20for%20matching%20EU%20social%20surveys15052012.pdf](http://www.q2012.gr/articlefiles/sessions/24.2%20Leulescu-Agafitei_A%20quality%20framework%20for%20matching%20EU%20social%20surveys15052012.pdf)
- (2013): *Statistical Matching: A model based approach for data integration*. Eurostat Methodologies and Working papers. Luxembourg, Publications Office of the European Union. Recuperado de: <http://ec.europa.eu/eurostat/documents/3888793/5855821/KS-RA-13-020-EN.PDF>
- LEULESCU, A., AGAFITEI, M., & MERCY, J.L.** (2011): *Statistical matching: a case study on EU-SILC and LFS. ESSnet Data Integration Workshop*. Madrid, noviembre de 2011. Recuperado de: www.ine.es/e/essnetdi_ws2011/.../Leulescu_Agafitei_Mercy.pdf

- LÓPEZ, A. (1993): "The Demand for Private Transport in Spain: a Microeconometric Approach". Working Paper 52, Universitat Pompeu Fabra.
- LÓPEZ-SALIDO, J. D. (1993): "Consumo y Ciclo Vital: Resultados para España con Datos de Panel". *Investigaciones Económicas*, Vol. 17, No.2, Mayo, pp. 285-312.
- LYDALL, H. F. (1979): *A Theory of Income Distribution*. London, Oxford University Press.
- MANKIW, G. (2010): *Macroeconomics*. 7th edition. New York, Worth Publishers.
- MASON, R. (2000): "The Social Significance of Consumption: James Duesenberry's Contribution to Consumer Theory," *Journal of Economic Issues*, 34, pp. 553-572. (<http://dx.doi.org/10.1080/00213624.2000.11506294>)
- MÉNDEZ, J.M. Y VEGA, P. (2011): *Linking data from administrative records and the Living Conditions Survey*. Working Paper 01/2011. INE.
- (2014): *Comparación de los ingresos del trabajo entre la Encuesta de Condiciones de Vida y las fuentes administrativas*. Working Paper 02/2014. INE.
- MEYER, B., WALLACE, M. & SULLIVAN, J. (2008): *The Under-Reporting of Transfers in Household Surveys: Its Nature and Consequences*. Working paper, University of Chicago. Recuperado de: <http://www.nber.org/programs/ag/rrc/08-12%20Meyer,%20Mok,%20Sullivan%20FINAL.pdf>
- MICHAEL, R.T. (1975): "Education and Consumption", in F. Thomas Juster, (Ed.), *Education, Income and Human Behavior*, pp. 235-252. New York: Columbia University Press for NBER. Recuperado de: <http://www.nber.org/chapters/c3699.pdf>
- MICKLEWRIGHT, J. AND SCHNEPF, S. V. (2010): "How reliable are income data collected with a single question". *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, Vol. 173, No. 2, pp. 409-429. (<http://dx.doi.org/10.1111/j.1467-985X.2009.00632.x>)
- MINCER, J. (1970): "The distribution of labor incomes: A survey with special references to the human capital approach". *Journal of Economic Literature*, Vol. 8, No. 1, pp. 1-27.
- (1974): *Schooling, experience and earnings*. New York, National Bureau of Economic Research.
- MODIGLIANI, F. (1971): "Monetary Policy and Consumption: linkages via interest rate and wealth effects in the FMP model", in FRB of Boston [eds.]: *Consumer Spending and Monetary Policy: the linkages*, pp. 9-84, FRB of Boston, Boston.
- MODIGLIANI, F. AND BRUMBERG, R. (1954): "Utility Analysis and the Consumption Function: An Interpretation of Cross-section Data" in Kenneth K. Kurihara, ed., *Post-Keynesian Economics*, New Brunswick, N.J., Rutgers University Press. pp. 388-436.

- MODIGLIANI, F. AND STEINDEL, CH. (1977): "Is a Tax Rebate an Effective Tool for Stabilization Policy?", *Brookings Papers on Economic Activity*, 1, pp. 177-203. (<http://dx.doi.org/10.2307/2534259>)
- MOORE, J.C, STINSON, L.L., & WELNIAK, E.J. (2000): "Income Measurement Error in Surveys: A review". *Journal of Official Statistics*, Vol. 66, No. 4, pp. 331-361. Recuperado de: <https://www.census.gov/srd/papers/pdf/sm97-05.pdf>
- MORENO, A. (dir.) (2003): *La distribución espacial de la renta en la Comunidad de Madrid. Análisis y aplicaciones*. Madrid, Ed. Comunidad de Madrid.
- MORENO, B. (2006): "Lugar de residencia y pautas de consumo. El Penedés y Barcelona, 1770-1790". *Revista de historia industrial*, 31, pp. 139-166.
- MUNNELL, A. H. AND CHEN, A. (2014): "Do census data understate retirement income?" *Issues in Brief* ib2014-19, Center for Retirement Research. Recuperado de: http://crr.bc.edu/wp-content/uploads/2014/12/IB_14-19-508.pdf
- MURCIA, A. (2007): "Determinantes del acceso al crédito de los hogares colombianos". *Ensayos Sobre Política Económica*, Vol. 25, No. 55, pp. 40-83. Recuperado de: <http://www.banrep.gov.co/sites/default/files/publicaciones/pdfs/borra449.pdf>
- NAVARRO, J.L. Y HERNÁNDEZ, E. (2004): "Distribución y redistribución de la renta en la literatura española reciente". *Estudios de Economía Aplicada*, Vol. 22, No. 1, pp. 29-65. Recuperado de: <https://dialnet.unirioja.es/descarga/articulo/858534.pdf>
- NORDBERG, L., PENTILÄ, I. & SANDSTRÖM, S. (2001): "Earnings Data from Surveys and Registers. Experiences from the Finnish European Community Household Panel Survey in 1996". *Paper prepared for the Intermediate Workshop of CHINTEX in Helsinki, 20-30 November 2001 - Statistics Finland*. Recuperado de: https://www.destatis.de/DE/Methoden/Methodenpapiere/Chintex/ResearchResults/IntermediateWorkshop/Downloads/WorkingPaper_wp5_sta.pdf;jsessionid=C954B862A489D23B73E655DFD5D1E87F.cae2?_blob=publicationFile
- NÚÑEZ, J. A. (2008): "Consumo y ahorro de los hogares jóvenes. Evidencia para España con datos de sección cruzada". *Cuadernos de Ciencias Económicas y Empresariales*, 54, pp. 57-81. Recuperado de: <http://cuadernos.uma.es/pdfs/pdf652.pdf>
- ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD) (2011): *How's life? Measuring Well-Being*. Paris, OECD Publishing.
- (2013a): *OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth*. Paris, OECD Publishing. Recuperado de: <http://www.oecd.org/statistics/302013041e.pdf>
- (2013b): *OECD Guidelines for Micro Statistics on Household Wealth*. Paris, OECD Publishing. Recuperado de: <http://www.oecd.org/statistics/OECD-Guidelines-for-Micro-Statistics-on-Household-Wealth.pdf>

- OLIVER, J., X. RAMOS, X. Y RAYMOND, J. L. (2001): "Anatomía de la distribución de la renta en España, 1985-1996: la continuidad de la mejora". *Papeles de Economía Española*, No. 88, pp. 67-88.
- ORTEGA, J.F. (2000): *Nuevas Familias de Estimadores Robustos y Detección de Observaciones Atípicas en Modelos Lineales*. [Tesis Doctoral]. Facultad de Ciencias Económicas y Empresariales de Albacete. Universidad de Castilla-La Mancha.
- ORTEGA, J.F. Y CALLEALTA, F.J. (2002): "Ocultación de los Ingresos personales en España: una aproximación mediante procedimientos robustos". *Estadística Española*, Vol. 44, No. 149, pp. 47-68.
- PATINKIN, D. (1951): *Dinero, Interés y Precios*. Madrid, Editorial Aguilar.
- (1956): *Money, Interest and Prices: an integration of monetary and value theory*. Evanston, Peterson.
- PASCUAL, M. Y SARABIA, J.M. (2001): "Factores determinantes de la distribución personal de la renta: Un estudio empírico a partir del PHOGUE". *Papeles de Trabajo* No. 15/04. Instituto de Estudios Fiscales. Recuperado de: http://www.ief.es/documentos/recursos/publicaciones/papeles_trabajo/2004_15.pdf
- (2003): "Modelización de la distribución personal de la renta en España: Un análisis regional". *Revista de Estudios Regionales*, 80, pp. 13-41. Recuperado de: <http://www.revistaestudiosregionales.com/documentos/articulos/pdf1019.pdf>
- (2004): "La distribución personal de la renta: Subdeclaración de ingresos y análisis dinámico a partir del PHOGUE". *XII Encuentro de Economía Pública*. Palma de Mallorca, 3 y 4 de febrero de 2005.
- (2006): "La distribución personal de la renta en España: Un estudio de las principales fuentes estadísticas". *Estadística Española*, Vol. 48, No. 163, pp. 401-432.
- PASHARDES, P. AND POLYCARPOU, A. (2008): "Income tax evasion, inequality and poverty". *Cyprus Economic Policy Review*, Vol. 2, No. 2, pp. 37-49. Recuperado de: https://ucy.ac.cy/erc/documents/Full_Text_Pashardes1208.pdf
- PAULUS, A. (2015): *Income underreporting based on income-expenditure gaps: survey vs tax records*. ISER Working Paper Series No. 2015-15. Institute for Social and Economic Research (ISER), University of Essex, Colchester, UK. Recuperado de: <https://www.iser.essex.ac.uk/research/publications/working-papers/iser/2015-15.pdf>
- PENA, B., CALLEALTA, F.J., CASAS, J.M., NÚÑEZ, J.J. Y MEREDIZ, A. (1996): *Distribución Personal de la Renta en España*. Madrid, Pirámide.
- PENA, B.; CALLEALTA, F.J. Y NÚÑEZ, J. (2000): "Encuestas de Presupuestos Familiares, Renta de las Familias y Estudio de la Distribución Personal de la Renta: Una experiencia española". *Conference of Consequences of Financial Crisis on*

- Income Distribution and Poverty in Latin America*, México D.F. (México), Mayo, 2000. Recuperado de: <http://ftp.itam.mx/pub/investigadores/delnegro/poverty.htm>
- PÉREZ, C., VILLANUEVA, J., BURGOS, M^a. J., BERMEJO, E. Y KHALIFI CHAIRI EL KAMMEL, L. (2016): *La Muestra de IRPF de 2013: Descripción General y Principales Magnitudes*. Documentos de Trabajo No. 9, 2016. Instituto de Estudios Fiscales. Recuperado de: http://www.ief.es/documentos/recursos/publicaciones/documentos_trabajo/2016_09.pdf
- PÉREZ, C., VILLANUEVA, J., BURGOS, M^a. J., PRADELL, E. Y GALLEGO, C. (2016): *Panel de declarantes del IRPF 1999-2012: metodología, estructura y variables*. Primera edición. Documentos de Trabajo No. 11/2016. Instituto de Estudios Fiscales. Recuperado de: http://www.ief.es/documentos/recursos/publicaciones/documentos_trabajo/2016_11.pdf
- PHELPS BROWN, H. (1977): *The Inequality of Pay*. Oxford, Oxford University Press.
- PICOS, F. (2003): *Las rentas del capital en PHOGUE: análisis de su fiabilidad y corrección mediante fusión estadística*. Papeles de Trabajo No. 31/03 Instituto de Estudios Fiscales. Madrid. Recuperado de: http://www.ief.es/documentos/recursos/publicaciones/papeles_trabajo/2003_31.pdf
- (2006): "Microsimulación mediante fusión de PHOGUE y panel de declarantes para evaluar reformas fiscales". *Revista de Economía Aplicada*, Vol. 41, No. 14, pp. 33-60. Recuperado de: www.revecap.com/revista/numeros/41/pdf/picos.pdf
- PIGOU, A.C. (1943): "The classical stationary state". *Economic Journal*, 53, pp. 343-351. (<http://dx.doi.org/10.2307/2226394>)
- PRIETO, J., SANZO, M.J., Y SUÁREZ, J. (2006): "Análisis económico de la actitud hacia el fraude fiscal en España". *Hacienda Pública Española*, No. 177, pp. 107-128. Recuperado de: http://www.ief.es/documentos/recursos/publicaciones/revistas/hac_pub/177_Fraude.pdf
- PRIETO, M. Y PENA, B. (2000): "Repercusiones de la ocultación de renta sobre la medición de la desigualdad". *Estudios de Economía Aplicada*, Vol 14, No. 1, pp. 153-172. Recuperado de: <https://dialnet.unirioja.es/descarga/articulo/1217233.pdf>
- RADNER, D.B, RICH, A., GONZALEZ M.E., JABINE, T.B. & MULLER, H.J. (1980): *Report on Exact and Statistical Matching Techniques*. Statistical Working Paper 5, Office of Federal Statistical Policy and Standards, US DoC.
- RAMOS, B. (2007): "Comparación de las magnitudes estadísticas del Empleo según la Encuesta de Población Activa y la Muestra Continua de Vidas Laborales". *Ponencia presentada en las Jornadas de Usuarios de la Muestra Continua de Vidas Laborales organizadas por FEDEA y la Tesorería General de la Seguridad Social, octubre de 2007, Madrid*.

- RASCHE, R.H. (1972): "Impact of the Stock Market on Private Demand". *American Economic Review*, 62, pp. 220-28.
- RASCHE, R. AND SHAPIRO, H. (1968): "The FRB-MIT Econometric Model: Its Special Features". *American Economic Review*, 58, pp. 123-49.
- RASTRIGINA, O., SUTHERLAND, H., LEVENTI, C., LOPEZ-VILAPLANNA, C. & LEULESCU, A. (2016): "'Flash estimates' of income distribution indicators for the European Union: Methods, assessment and future prospects". *International Association for Research in Income and Wealth*. Dresden. 2016. Conference paper.
- RAVALLION, M. (2001): *Measuring Aggregate Welfare in Developing Countries: How Well do National Accounts and Surveys Agree?* Policy Research Working Paper 2665, World Bank. (<http://dx.doi.org/10.1596/1813-9450-2665>)
- REGLAMENTO (CE) No. 2223/96 del Consejo de 25 de junio de 1996 relativo al Sistema Europeo de Cuentas Nacionales y Regionales de la Comunidad. Publicado en el Diario Oficial de la Unión Europea nº L 310 de 30/11/1996, pp. 0001 - 0469. Recuperado de: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:1996R2223:20071230:ES:PDF>
- No. 1177/2003 del Parlamento Europeo y del Consejo de 16 de junio de 2003, relativo a las estadísticas comunitarias sobre la renta y las condiciones de vida (EU-SILC). Publicado en el Diario Oficial de la Unión Europea nº L 165 de 03/07/2003, pp. 0001 - 0009. Recuperado de: <http://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32003R1177&from=ES>
- REGLAMENTO (UE) No. 549/2013 del Parlamento Europeo y del Consejo, de 21 de mayo de 2013, relativo al Sistema Europeo de Cuentas Nacionales y Regionales de la Unión Europea. Publicado en el Diario Oficial de la Unión Europea nº L 174 de 26/06/2013, pp. 0001 - 0727. Recuperado de: <http://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32013R0549&from=ES>
- RODGERS, W. L. (1984): An Evaluation of Statistical Matching. *Journal of Business & Economic Statistics*, Vol. 2, No. 1, pp. 91-102. (<http://dx.doi.org/10.2307/1391358>)
- ROMER, C. (1990): "The Great Crash and the Onset of the Great Depression". *Quarterly Journal of Economics*, 422(3), pp. 597-624. (<http://dx.doi.org/10.2307/2937892>)
- ROMER, D. (2012): *Advanced Macroeconomics, fourth edition*. New York, McGraw Hill.
- RUGGLES, N., AND RUGGLES, R. A. (1974): "A Strategy for Merging and Matching Microdata Sets". *Annals of Economic and Social Measurement*, Vol. 3, No. 2, pp. 353-371.
- RUÍZ-CASTILLO, J. (1987): *La medición de la pobreza y la desigualdad en España 1980-1981*. Servicio de Estudios del Banco de España. Estudios Económicos, 42. Banco de España, Madrid.

- RUIZ-HUERTA, J. Y MARTÍNEZ, R.** (1994): "La pobreza en España: ¿Qué nos muestran las encuestas de presupuestos familiares?". *Documentación Social*, 96, pp. 15-109.
- SAMUELSON, P.A.** (1941): "A statistical analysis of the consumption function". Appendix to Chapter 11 in A. H. Hansen (Ed.), *Fiscal Policy and Business Cycles*, pp. 250-260. New York, W. W. Norton.
- SANZ, M.B.** (1995): *La articulación micro-macro en el sector hogares: de la Encuesta de Presupuestos Familiares a la Contabilidad Nacional*. Papeles de Trabajo No. 27/95. Instituto de Estudios Fiscales. Madrid.
- SEN, A.** (2001): *El nivel de vida*. Madrid, Editorial Complutense.
- SMITH, R. S.** (1989): *Factors affecting saving, policy tools, and tax reform: A review*. IMF Working Paper Series IMF/89/47. Washington, International Monetary Fund. (<http://dx.doi.org/10.2307/3867303>)
- SOSA, R. M.** (2009): *Wealth Effects on Consumption: Evidence from the Euro Area*. Working Paper Series No. 1050. European Central Bank. Frankfurt am Main. Recuperado de: <https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1050.pdf>
- STIGLITZ, J.E., SEN, A. & FITOUSSI, J.P.** (2009): *Report by the Commission on the Measurement of Economic Performance and Social Progress*. Paris, Commission on the Measurement of Economic Performance and Social Progress. Recuperado de: <http://ec.europa.eu/eurostat/documents/118025/118123/Fitoussi+Commission+report/7bac2480-4658-439f-b022-e6542ebf714e>
- SUSIN, S.** (2003): *Discrepancies Between Measured Income in the American Housing Survey (AHS) and the Current Population Survey (CPS): Final Report*. U.S. Census Bureau Staff Paper on Income. No. 2003/04. Washington, DC.
- SUTHERLAND, H., TAYLOR, R. & GOMULKA J.** (2001): *Combining household income and expenditure data in policy simulations*. University of Cambridge Department of Applied Economics Working Papers / Microsimulation Unit Discussion Papers, No. MU0101. (<http://dx.doi.org/10.1111/1475-4991.00066>)
- TAUBMAN, P.** (1975): *Sources of Inequality of Earnings*. Amsterdam, North-Holland.
- (1977): "The Relative Influence of Inheritable and Environmental Factors and the Importance of Intelligence in Earnings Functions". *IEA Conference Paper, Nordwijk*.
- TEMIN, P.** (1976): *Did Monetary Forces Cause the Great Depression?*. New York, Norton.
- TINBERGEN, J.** (1975): *Income Distribution: Analysis and Policies*. Amsterdam, Oxford University Press.
- TROITIÑO, A.** (2001): "Incidencia de la ocultación de ingresos de la EPF sobre los niveles regionales de desigualdad relativa y bienestar-renta". *Estudios de Economía Aplicada*, Vol. 19, No. 3, pp. 189-210. Recuperado de: <https://dialnet.unirioja.es/descarga/articulo/1218587.pdf>

- UNITED NATIONS (UN)** (1993): *System of National Accounts 1993*. New York, United Nations.
- (1999): *Handbook of Input-Output Table Compilation and Analysis*. Studies in Methods, Series F, No. 74. New York, United Nations . Recuperado de: https://unstats.un.org/unsd/publication/SeriesF/SeriesF_74E.pdf
- UNITED NATIONS, EUROPEAN COMMISSION, INTERNATIONAL MONETARY FUND, ORGANISATION FOR ECONOMIC CO-OPERATION & DEVELOPMENT AND WORLD BANK** (2009): *System of National Accounts 2008*. Sales No. E.08.XVII.29. Recuperado de: <http://unstats.un.org/unsd/nationalaccount/docs/SNA2008.pdf>
- UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE (UNECE)** (2011): *The Canberra Group Handbook on Household Income Statistics, Second Edition*. Ginebra, United Nations.
- WANG, X. AND WOO, W. T.** (2011): "The Size and Distribution of Hidden Household Income in China". *Asian Economic Papers*, Vol. 10, No. 1, pp. 1-26. (http://dx.doi.org/10.1162/ASEP_a_00064)
- WARD, M. D.** (1978): *The Political Economy of Distribution: Equality Versus Inequality*. New York: Elsevier.
- WRIGHT, E.O.** (1979): *Class Structure and Income Determination*. New York, Academic Press.
- ZABALZA, A. Y ANDRÉS, J.** (1991): "¿Afecta la fiscalidad al ahorro?". *Moneda y Crédito*, 192, pp. 41-75.
- ZELLNER, A.** (1957): "The short-run consumption function". *Econometrica*, Vol. 25, No. 4, pp. 552-567. (<http://dx.doi.org/10.2307/1905383>)

Anexos

Anexo 1. Programas informáticos para la construcción de variables homogéneas en los procesos de fusión de la ECV-EPF y ECV-EPF-EFF

Anexo 1.a. Programa que prepara los Datos de Personas de las EPF para matching

```

=====
*====>>> 03a_Prepara_Datos_EPF_P_para_matching.sps
=====
.
* Funde los archivos de SPSS para tomar las variables necesarias y crea un archivo con la estructura de la BDH.

define !Prepara_Datos_EPF_P_para_matching (epf = !TOKENS(1) ).

New file.
dataset close all.

OUTPUT NEW NAME =salida.
OUTPUT ACTIVATE salida.

CD 'WorkDir'.

=====
* Carga de la BDP en memoria .
=====

GET FILE='BDPDir\BDP.sav'.
DATASET NAME BDP.

=====
* Carga de los datos básicos de las personas, P , de la EPF.
=====

!!let !arch=!quote(!concat('FuentesDir\EPF\' , !epf, '\FichusP',!epf, '.sav')).
GET FILE=!arch.
DATASET NAME ArchivoP.

*----- idp: Identificador de la Persona -----

DATASET ACTIVATE ArchivoP .
compute idp=100*NUMERO + NORDEN .
execute.
APPLY DICTIONARY
  /FROM BDP /SOURCE VARIABLES=idp
  /TARGET VARIABLES=idp
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

=====
* Preparación y creación de algunas variables auxiliares para las personas .
=====

*--- idh y num: cálculo del identificador de su hogar y del número de orden de la persona dentro del hogar.

DATASET ACTIVATE ArchivoP .
COMPUTE idh = NUMERO .
compute num=NORDEN .
execute.
APPLY DICTIONARY
  /FROM BDP /SOURCE VARIABLES=idh
  /TARGET VARIABLES=idh
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
APPLY DICTIONARY
  /FROM BDP /SOURCE VARIABLES=num
  /TARGET VARIABLES=num
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*--- wper y wper16: cálculo de pesos de todeas las personas y para estudios sólo de adultos.
* Ojo: hay personas con ingresos, que se han tenido en cuenta a la hora de calcular la renta del hogar, que tienen peso
PB040=0 en el fichero P de personas.

DATASET ACTIVATE ArchivoP .
compute wper=FACTOR.
compute wper16=$SYSMIS.

```

```

if (EDAD>=16) wper16=FACTOR.
execute.
APPLY DICTIONARY
  /FROM BDP /SOURCE VARIABLES=wper
  /TARGET VARIABLES=wper
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
APPLY DICTIONARY
  /FROM BDP /SOURCE VARIABLES=wper16
  /TARGET VARIABLES=wper16
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*=====
*==== Cálculo de principales variables sociodemográficas de la persona .
*=====

*----- sexop: Sexo de la Persona -----.

RECODE SEXO (MISSING=SYSMIS) (-9=SYSMIS) (1=1) (6=6) (ELSE=Copy) INTO sexop.
execute.
APPLY DICTIONARY
  /FROM BDP /SOURCE VARIABLES=sexop
  /TARGET VARIABLES=sexop
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
CROSSTABS SEXO by sexop.

*----- edadp: Edad de la Persona (a 1 de enero) -----.

RECODE EDAD (-9 MISSING=SYSMIS) (79 THRU HIGHEST=79) (ELSE=Copy) INTO edadp.
execute.
APPLY DICTIONARY
  /FROM BDP /SOURCE VARIABLES=edadp
  /TARGET VARIABLES=edadp
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
CROSSTABS EDAD by edadp.

*----- gedadp: Grupo de Edad de la Persona -----.

compute gedadp=$SYSMIS.
if (edadp ge 0 and edadp le 15) gedadp=0 .
if (edadp gt 15 and edadp le 25) gedadp=1 .
if (edadp gt 25 and edadp le 35) gedadp=2 .
if (edadp gt 35 and edadp le 45) gedadp=3 .
if (edadp gt 45 and edadp le 55) gedadp=4 .
if (edadp gt 55 and edadp le 65) gedadp=5 .
if (edadp gt 65) gedadp=6 .
execute.
APPLY DICTIONARY
  /FROM BDP /SOURCE VARIABLES=gedadp
  /TARGET VARIABLES=gedadp
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*----- Cumplidos14: Variables indicadora de si la persona tiene 14 años o mas .

do if (Missing(edadp)).
  compute Cumplidos14=$Sysmis.
else if (edadp >= 14).
  compute Cumplidos14=1.
else if (0<=edadp< 14).
  compute Cumplidos14=0.
end if.
EXECUTE.
APPLY DICTIONARY
  /FROM BDP /SOURCE VARIABLES=Cumplidos14
  /TARGET VARIABLES=Cumplidos14
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*----- Cumplidos16: Variables indicadora de si la persona tiene 16 años o mas .

do if (Missing(edadp)).
  compute Cumplidos16=$Sysmis.
else if (edadp >= 16).
  compute Cumplidos16=1.
else if (0<=edadp and edadp< 16).
  compute Cumplidos16=0.
end if.
EXECUTE.
APPLY DICTIONARY
  /FROM BDP /SOURCE VARIABLES=Cumplidos16
  /TARGET VARIABLES=Cumplidos16
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

```

* ----- Cumplidos18: Variables indicadora de si la persona tiene 18 años o mas .

```
do if (Missing(edadp)).
  compute Cumplidos18=$Sysmis.
else if (edadp >= 18).
  compute Cumplidos18=1.
else if (0<=edadp and edadp< 18).
  compute Cumplidos18=0.
end if.
EXECUTE.
APPLY DICTIONARY
  /FROM BDP /SOURCE VARIABLES=Cumplidos18
  /TARGET VARIABLES=Cumplidos18
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
```

*---- estcivlp: Estado Civil de la Persona -----.

```
!if (!epf !le 'EPF10' ) !then
RECODE ECIVIL (-9 MISSING=SYSMIS) (1=1) (2=2) (3=4) (4=3) (5=5) (ELSE=Copy) INTO estcivlp.
!lifend
!if (!epf !ge 'EPF11' ) !then
RECODE ECIVILLEGAL (-9 MISSING=SYSMIS) (1=1) (2=2) (3=4) (4=3) (5=5) (ELSE=Copy) INTO estcivlp.
!lifend
if (gedadp=0) estcivlp=0.
EXECUTE.
APPLY DICTIONARY
  /FROM BDP /SOURCE VARIABLES=estcivlp
  /TARGET VARIABLES=estcivlp
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
!if (!epf !le 'EPF10' ) !then
CROSSTABS ECIVIL by estcivlp.
!lifend
!if (!epf !ge 'EPF11' ) !then
CROSSTABS ECIVILLEGAL by estcivlp.
!lifend
```

*----- nedp: Nivel de Estudios de la Persona -----.

```
RECODE ESTUDRED (-9 MISSING=SYSMIS) (1=2) (2=3) (3=4) (4=5) INTO nedp.
if (gedadp=0) nedp=0.
execute.
if ( ESTUDIOS=1 and not ESTUDRED>1) nedp=1 .
execute.
APPLY DICTIONARY
  /FROM BDP /SOURCE VARIABLES=nedp
  /TARGET VARIABLES=nedp
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
CROSSTABS ESTUDRED by nedp.
CROSSTABS ESTUDIOS by nedp.
```

*----- nacip: Nacionalidad de la Persona -----.

```
RECODE NACIONA (-9 MISSING=SYSMIS) (1 3=1) (2=2) INTO nacip.
do if (NACIONA=2).
!if (!epf !le 'EPF10' ) !then
  if (PAIS=1)          nacip=2.
  if (PAIS=2 or PAIS=3)  nacip=3.
!lifend
!if (!epf !ge 'EPF11' ) !then
  if (PAISNACION=1)  nacip=2.
  if (PAISNACION=2 or PAISNACION=3)  nacip=3.
!lifend
end if.
if (gedadp=0) nacip=0.
execute.
APPLY DICTIONARY
  /FROM BDP /SOURCE VARIABLES=nacip
  /TARGET VARIABLES=nacip
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
```

*----- situac1p: situación de la persona en relación con la actividad (1ª calificación) -----.

```
compute situac1p=$SYSMIS.
if (gedadp=0) situac1p=0.
execute.
APPLY DICTIONARY
  /FROM BDP /SOURCE VARIABLES=situac1p
  /TARGET VARIABLES=situac1p
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
```

*----- situac2p: situación de la persona en relación con la actividad (2ª calif) -----.

```
compute situac2p=$SYSMIS.
if (gedadp=0) situac2p=0.
execute.
APPLY DICTIONARY
  /FROM BDP /SOURCE VARIABLES=situac2p
  /TARGET VARIABLES=situac2p
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
```

*----- situac3p: situación de la persona en relación con la actividad (3ª calif) -----.

```
RECODE SITUACT (-9 MISSING=SYSMIS) (1 2=1) (3=2) (4=4) (5=3) (6=5) (7=6) INTO situac3p.
if (gedadp=0) situac3p=0.
execute.
APPLY DICTIONARY
  /FROM BDP /SOURCE VARIABLES=situac3p
  /TARGET VARIABLES=situac3p
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
CROSSTABS SITUACT by situac3p.
CROSSTABS SITUACT by situac3p.
```

*----- situatp: situación profesional de la persona (1ª clasif) -----.

```
compute situatp=$SYSMIS.
if (gedadp=0) situatp=0.
execute.
APPLY DICTIONARY
  /FROM BDP /SOURCE VARIABLES=situatp
  /TARGET VARIABLES=situatp
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
```

*----- ocupacp: Ocupación laboral de la persona -----.

```
compute ocupacp=$SYSMIS.
if (gedadp=0) ocupacp=0.
execute.
APPLY DICTIONARY
  /FROM BDP /SOURCE VARIABLES=ocupacp
  /TARGET VARIABLES=ocupacp
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
```

*----- csep: Categoría socio-económica de la Persona -----.

```
compute csep=$SYSMIS.
if (gedadp=0) csep=0.
execute.
APPLY DICTIONARY
  /FROM BDP /SOURCE VARIABLES=csep
  /TARGET VARIABLES=csep
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
```

*---- activestb1p: Actividad del establecimiento laboral (1ª clasif) -----.

```
STRING activestb1p (A5).
compute activestb1p = ' '.
if (gedadp=0) activestb1p = '_ND'.
execute.
APPLY DICTIONARY
  /FROM BDP /SOURCE VARIABLES=activestb1p
  /TARGET VARIABLES=activestb1p
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
```

*---- activestb2p: Actividad del establecimiento laboral (2ª clasif) -----.

```
STRING activestb2p (A5).
compute activestb2p = ' '.
if (gedadp=0) activestb2p = '_ND'.
execute.
APPLY DICTIONARY
  /FROM BDP /SOURCE VARIABLES=activestb2p
  /TARGET VARIABLES=activestb2p
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
```

*---- tipcontp: Tipo de contrato de la persona -----.

```
compute tipcontp=$SYSMIS.
if (gedadp=0) tipcontp=0.
execute.
APPLY DICTIONARY
  /FROM BDP /SOURCE VARIABLES=tipcontp
```

```
/TARGET VARIABLES=tipcontp
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
```

*---- dep: indicador de si la persona es dependiente o no, según la ECV -----.

```
compute dep=$SYSMIS.
do if (edadp<18).
  compute dep=1.
else if ( (edadp <25) and (3<=situac3p and situac3p<=6) ) .
  compute dep=1.
else if ( (edadp >=25) or (1<=situac3p and situac3p<=2) ).
  compute dep=0.
end if.
execute.
APPLY DICTIONARY
/FROM BDP /SOURCE VARIABLES=dep
/TARGET VARIABLES=dep
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
```

*---- ocup: indicador de si la persona esta ocupada o no -----.

```
compute ocup=$SYSMIS.
do if (edadp<16).
  compute ocup=0.
else if ( situac3p=1 ).
  compute ocup=1.
else if ( situac3p>1 ).
  compute ocup=0.
end if.
execute.
APPLY DICTIONARY
/FROM BDP /SOURCE VARIABLES=ocup
/TARGET VARIABLES=ocup
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
```

*---- dep16: indicador de si la persona es dependiente o no, según la EPF -----.

```
compute dep16=$SYSMIS.
do if (edadp<16).
  compute dep16=1.
else if ( (edadp <25) and (3<=situac3p and situac3p<=6) ) .
  compute dep16=1.
else if ( (edadp >=25) or (1<=situac3p and situac3p<=2) ).
  compute dep16=0.
end if.
execute.
APPLY DICTIONARY
/FROM BDP /SOURCE VARIABLES=dep16
/TARGET VARIABLES=dep16
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
```

*** comprobación -----.

CROSSTABS ninodep by dep16.

*---- adul: indicador de si la persona es adulta (16 o más años) -----.

```
compute adul=$SYSMIS.
do if (edadp<16) .
  compute adul=0.
else if (edadp >=16).
  compute adul=1 .
end if.
execute.
APPLY DICTIONARY
/FROM BDP /SOURCE VARIABLES=adul
/TARGET VARIABLES=adul
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
```

*---- adulnodep: indicador de si la persona tiene 18 o más años y es no dependiente -----.

```
compute adulnodep=$SYSMIS.
do if (Cumplidos18 and not dep) .
  compute adulnodep=1.
else if (not Cumplidos18 or dep) .
  compute adulnodep=0 .
end if.
execute.
APPLY DICTIONARY
/FROM BDP /SOURCE VARIABLES=adulnodep
/TARGET VARIABLES=adulnodep
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
```

```

*=====
*==== Datos de Ingresos de la persona =====
*=====

If (interinp=-9) interinp=$SYSMIS.
if (impexacp=-9) impexacp=$SYSMIS.
do if (percep=6 ).
    compute interinp=0.
    compute impexacp=0.
end if.
execute.

*----- Ing_Np: Total de Ingresos personales netos -----

compute Ing_Np= IMPEXACP*12.
execute.
formats Ing_Np (F10.2).
variable labels Ing_Np 'Total de Ingresos Personales Netos'.
RECODE
Ing_Np
(MISSING=SYSMIS) (LOWEST thru -0,001=0) (0 thru 5999,999=1) (6000 thru 11999,999=2) (12000 thru 17999,999=3) (18000 thru
23999,999=4)
(24000 thru 29999,999=5) (30000 thru 35999,999=6) (36000 thru 59999,999=7) (60000 thru 83999,999=8) (84000 thru
107999,999=9)
(108000 thru HIGHEST =10) INTO IN_Ing_Np.
execute.
*----- comprobación -----
CROSSTABS IN_Ing_Np by INTERINP.

*---- perc: indicador de si la persona percibe ingresos o no -----

do if ( missing(interinp) ).
    compute perc = $sysmis.
else if ( interinp = 0 ).
    compute perc =0.
else.
    compute perc = 1.
end if.
execute.
formats perc (F1.0).
VARIABLE LABELS perc '¿Percibe ingresos?'.
Value Labels perc
1 'Sí'
0 'No'.
*----- comprobación -----
CROSSTABS perc by PERCEP.

*=====
*==== Cálculo del número de orden de la persona dentro del hogar, en función de su aportación económica =====
*=====

DATASET ACTIVATE ArchivoP.
compute casado=(estcivlp=2).
execute.
AGGREGATE
/OUTFILE=* MODE=ADDVARIABLES
/BREAK=idh
/n_adulnodep=SUM(adulnodep)
/n_casados=SUM(casado).

*----- caso de hogares con un adulto no dependientes -----
* (primera etapa) se ordenan los adultos no dependientes .
compute filtro=(adulnodep =1).
if missing(filtro) filtro=0.
execute.
SORT CASES BY idh(A) filtro (D) interinp (D) Ing_Np (D) edadp(D) sexop(A) num (A) .
execute.
do if (n_adulnodep=1) .
    DO IF ($casenum=1 or idh-lag(idh) ~= 0) .
        compute norden_ing=1.
    ELSE .
        compute norden_ing=1+lag(norden_ing).
    END IF.
end if.
execute.
DELETE VARIABLES filtro .

*----- caso de otros hogares -----
* (primera etapa) se ordenan los adultos no dependientes, con ingresos positivos, dando prioridad a los no solteros .

```

```

do if n_casados>=2.
  recode estcivlp (2 3 4 5=1) (1 =2) (else=SYSMIS) into orden.
else.
  compute orden=1.
end if.
compute filtro=(adulnodep =1 and interinp > 0 ).
if missing(filtro) filtro=0.
execute.
SORT CASES BY idh(A) filtro (D) orden(A) interinp (D) lng_Np (D) edadp(D) sexop(A) num (A) .
execute.
do if not (n_adulnodep=1) .
  DO IF ($casenum=1 or idh-lag(idh) ~= 0) .
    compute norden_ing=1.
  ELSE .
    compute norden_ing=1+lag(norden_ing).
  END IF.
  if (not filtro) and not missing(norden_ing) norden_ing=999999999.
end if.
execute.
* (segunda etapa) se ordenan el resto de adultos no dependientes activos, y luego el resto de miembros .
compute filtro=(filtro) or (adulnodep =1 and (situac3p=1 or situac3p=2)) .
if missing(filtro) filtro=0.
execute.
SORT CASES BY idh(A) filtro (D) norden_ing (A) interinp (D) lng_Np (D) sexop(A) num (A) .
execute.
do if not (n_adulnodep=1).
  DO IF ($casenum=1 or idh-lag(idh) ~= 0).
    compute norden_ing=1.
  ELSE .
    compute norden_ing=1+lag(norden_ing).
  END IF.
end if.
execute.
formats norden_ing (F2.0).
variable labels norden_ing 'número de orden en importancia económica dentro del hogar'.
DELETE VARIABLES filtro orden n_casados n_adulnodep .
SORT CASES BY idp(A) .

=====
* Grabación de Datos.
=====

SORT CASES BY idh (A) num(A).
EXECUTE .

!!let !arch=!quote(!concat('ArchivoP_', !pf, '_para_matching.sav')).
save OUTFILE=!arch /COMPRESSED .

=====
* Salida .
=====

!!let !arch=!quote(!concat('ArchivoP_', !pf, '_para_matching.spv')).
OUTPUT SAVE NAME=salida OUTFILE=!arch LOCK=NO.
OUTPUT CLOSE salida.

New file.
dataset close all.

lenddefine.

```

Anexo 1.b. Programa que prepara los Datos de Hogares de las EPF para matching

```

=====
*==>>> 03a_Genera_Datos_EPF_H_para_matching.sps
=====
.
* Funde los archivos de SPSS para tomar las variables necesarias y crea un archivo con la estructura de la BDH.

define !Prepara_Datos_EPF_H_para_matching (epf = !TOKENS(1) ).

New file.
dataset close all.

OUTPUT NEW NAME =salida.
OUTPUT ACTIVATE salida.

CD 'WorkDir'.

=====
* Carga de la BDH_Match en memoria .
=====

GET FILE='F_BDHDDir\BDH_para_matching.sav'.
DATASET NAME BDH.

=====
* Carga del archivo de personas en Memoria
=====

!let !arch=!quote(!concat('ArchivoP_', !epf, '_para_matching.sav')).
GET FILE=!arch.
DATASET NAME ArchivoP.

=====
* Carga de los datos básicos del Hogar, H, de la EPF .
=====

!let !arch=!quote(!concat('FuentesDir\EPF\'', !epf, '\FichusH', !epf, '.sav')).
GET FILE=!arch.
DATASET NAME ArchivoH.

*----- Pais: Pais del Hogar -----,

DATASET ACTIVATE ArchivoH.
compute Pais=34.
execute.
APPLY DICTIONARY
  /FROM BDH /SOURCE VARIABLES=Pais
  /TARGET VARIABLES=Pais
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*----- ano: año de referencia de los gastos -----,

DATASET ACTIVATE ArchivoH.
compute ano=ANOENC.
execute.
APPLY DICTIONARY
  /FROM BDH /SOURCE VARIABLES=ano
  /TARGET VARIABLES=ano
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*----- idh: Identificador del Hogar -----,

DATASET ACTIVATE ArchivoH.
compute idh=NUMERO.
execute.
APPLY DICTIONARY
  /FROM BDH /SOURCE VARIABLES=idh
  /TARGET VARIABLES=idh
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*----- w: Factor de elevación del hogar -----,

DATASET ACTIVATE ArchivoH .
compute w=FACTOR.
APPLY DICTIONARY
  /FROM BDH /SOURCE VARIABLES=w

```



```

/TARGET VARIABLES=w
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
*=====
* Composicion y Tipología del Hogar .
*=====

* ----- nmh: Número de Miembros del Hogar -----

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
AGGREGATE
/OUTFILE=Aux
/BREAK=idh
/nmh = N.
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
/Table=Aux
/BY idh.
execute.
APPLY DICTIONARY
/FROM BDH /SOURCE VARIABLES=nmh
/TARGET VARIABLES=nmh
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
*----- comprobación -----
CROSSTABS nmh by NMIEMB.

*----- nmh16: Número de miembros del Hogar con 16 años o más -----

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
AGGREGATE
/OUTFILE=Aux
/BREAK=idh
/nmh16 = SUM(Cumplidos16).
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
/Table=Aux
/BY idh.
execute.
APPLY DICTIONARY
/FROM BDH /SOURCE VARIABLES=nmh16
/TARGET VARIABLES=nmh16
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
*----- comprobación -----
CROSSTABS nmh16 by NMIEMB4.

*----- nmh14: Número de miembros del Hogar con 14 años o más -----

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
AGGREGATE
/OUTFILE=Aux
/BREAK=idh
/nmh14 = SUM( Cumplidos14 ) .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
/Table=Aux
/BY idh .
execute.
APPLY DICTIONARY
/FROM BDH /SOURCE VARIABLES=nmh14
/TARGET VARIABLES=nmh14
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
*----- comprobación -----
CROSSTABS nmh14 by NMIEMB1.

*=====
* Variables específicas del Hogar para el matching.
*=====

*----- MATCH - REGION_m: Comunidad Autónoma-----

DATASET ACTIVATE ArchivoH .
RECODE CCAA (MISSING=SYSMIS) (18 =20) (ELSE=Copy) INTO REGION_m.
execute.
APPLY DICTIONARY
/FROM BDH /SOURCE VARIABLES=REGION_m
/TARGET VARIABLES=REGION_m

```

```

/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
*--- comprobación ----.
RECODE CCAA (SYSMIS=-9) (ELSE=Copy) INTO filas.
RECODE REGION_m (SYSMIS=-9) (ELSE=Copy) INTO columnas.
execute.
APPLY DICTIONARY
/FROM ArchivoH /SOURCE VARIABLES=CCAA
/TARGET VARIABLES=filas
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
APPLY DICTIONARY
/FROM ArchivoH /SOURCE VARIABLES=REGION_m
/TARGET VARIABLES=columnas
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
CROSSTABS /TABLES=filas BY columnas /MISSING=INCLUDE.
DELETE VARIABLES filas columnas.

```

*----- MATCH - HAB_m : Habitat de la vivienda -----.

```

DATASET ACTIVATE ArchivoH .
RECODE DENSI (MISSING=SYSMIS) (1=1) (2=2) (3=3) (ELSE=Copy) INTO HAB_m.
execute.
APPLY DICTIONARY
/FROM BDH /SOURCE VARIABLES=HAB_m
/TARGET VARIABLES=HAB_m
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
*--- comprobación ----.
RECODE DENSI (SYSMIS=-9) (ELSE=Copy) INTO filas.
RECODE HAB_m (SYSMIS=-9) (ELSE=Copy) INTO columnas.
execute.
APPLY DICTIONARY
/FROM ArchivoH /SOURCE VARIABLES=DENSI
/TARGET VARIABLES=filas
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
APPLY DICTIONARY
/FROM ArchivoH /SOURCE VARIABLES=HAB_m
/TARGET VARIABLES=columnas
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
CROSSTABS /TABLES=filas BY columnas /MISSING=INCLUDE.
DELETE VARIABLES filas columnas.

```

*----- MATCH - TIPOVIV_m: Tipo de vivienda -----.

```

DATASET ACTIVATE ArchivoH .
RECODE TIPOEDIF (1=1) (2=2) (3=3) (4=4) (5=SYSMIS) (-9=SYSMIS) INTO TIPOVIV_m.
EXECUTE.
APPLY DICTIONARY
/FROM BDH /SOURCE VARIABLES=TIPOVIV_m
/TARGET VARIABLES=TIPOVIV_m
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
*--- comprobación ----.
RECODE TIPOEDIF (SYSMIS=-9) (ELSE=Copy) INTO filas.
RECODE TIPOVIV_m (SYSMIS=-9) (ELSE=Copy) INTO columnas.
execute.
APPLY DICTIONARY
/FROM ArchivoH /SOURCE VARIABLES=TIPOEDIF
/TARGET VARIABLES=filas
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
APPLY DICTIONARY
/FROM ArchivoH /SOURCE VARIABLES=TIPOVIV_m
/TARGET VARIABLES=columnas
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
CROSSTABS /TABLES=filas BY columnas /MISSING=INCLUDE.
DELETE VARIABLES filas columnas.

```

*----- MATCH - REGTENEVIV_m: Régimen de tenencia de la vivienda -----.

```

DATASET ACTIVATE ArchivoH .
RECODE REGTEN (1 2=1) (3=2) (4 5=3) (6=4) (MISSING=SYSMIS) INTO REGTENEVIV_m.
EXECUTE.
APPLY DICTIONARY
/FROM BDH /SOURCE VARIABLES=REGTENEVIV_m
/TARGET VARIABLES=REGTENEVIV_m
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
*--- comprobación ----.
RECODE REGTEN (SYSMIS=-9) (ELSE=Copy) INTO filas.
RECODE REGTENEVIV_m (SYSMIS=-9) (ELSE=Copy) INTO columnas.
execute.
APPLY DICTIONARY
/FROM ArchivoH /SOURCE VARIABLES=REGTEN
/TARGET VARIABLES=filas
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

```

```

APPLY DICTIONARY
/FROM ArchivoH /SOURCE VARIABLES=REGTENEVIV_m
/TARGET VARIABLES=columns
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
CROSSTABS /TABLES=filas BY columns /MISSING=INCLUDE.
DELETE VARIABLES filas columnas.

*----- MATCH - NUMHABIT_m: Número de habitaciones de la vivienda -----.

DATASET ACTIVATE ArchivoH .
RECODE NHABIT (1=1) (2=2) (3=3) (4=4) (5=5) (6 7 8=6) (-9=SYSMIS) INTO NUMHABIT_m.
EXECUTE.
APPLY DICTIONARY
/FROM BDH /SOURCE VARIABLES=NUMHABIT_m
/TARGET VARIABLES=NUMHABIT_m
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
*--- comprobación ---.
RECODE NHABIT (SYSMIS=-9) (ELSE=Copy) INTO filas.
RECODE NUMHABIT_m (SYSMIS=-9) (ELSE=Copy) INTO columnas.
execute.
APPLY DICTIONARY
/FROM ArchivoH /SOURCE VARIABLES=NHABIT
/TARGET VARIABLES=filas
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
APPLY DICTIONARY
/FROM ArchivoH /SOURCE VARIABLES=NUMHABIT_m
/TARGET VARIABLES=columns
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
CROSSTABS /TABLES=filas BY columnas /MISSING=INCLUDE.
DELETE VARIABLES filas columnas.

*----- MATCH - TIPHOGAR_m: Tipo de hogar -----.

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
AGGREGATE
/OUTFILE='aux'
/BREAK=idh
/max_edad=MAX(edadp)
/n_adult=SUM(adulnodep)
/n_dep=SUM(dep)
/n_desc=NMISS(dep).
DATASET ACTIVATE ArchivoH.
MATCH FILES /FILE=*
/TABLE='aux'
/BY idh.
EXECUTE.
dataset close aux.

DATASET ACTIVATE ArchivoH.
compute TIPHOGAR_m=$SYSMIS.
DO if (nmh=1 and n_desc=0 and n_adult=1).
  if (max_edad < 30) TIPHOGAR_m=1.
  if (30<=max_edad and max_edad<65) TIPHOGAR_m=2.
  if (max_edad >=65) TIPHOGAR_m=3.
ELSE if (nmh=1 and n_desc=0 and n_dep=1).
  compute TIPHOGAR_m=14.
ELSE IF (nmh=2) and (n_desc=0) .
  if (n_adult=2 and n_dep=0 and max_edad>=65) TIPHOGAR_m=7.
  if (n_adult=2 and n_dep=0 and max_edad <65) TIPHOGAR_m=8.
  if (n_adult=1) and (n_dep=1) TIPHOGAR_m=10.
  if (n_adult=0) and (n_dep=2) TIPHOGAR_m=14.
ELSE IF (nmh>2) and (n_desc=0).
  do if (n_dep=0).
    compute TIPHOGAR_m=9.
  else if (n_dep>0) .
    if (n_adult=0) TIPHOGAR_m=14.
    if (n_adult=1) TIPHOGAR_m=10.
    if (n_adult=2) and (n_dep=1) TIPHOGAR_m=11.
    if (n_adult=2) and (n_dep=2) TIPHOGAR_m=12.
    if (n_adult=2) and (n_dep>=3) TIPHOGAR_m=13.
    if (n_adult>2) TIPHOGAR_m=14.
  end if.
END IF.
EXECUTE.
APPLY DICTIONARY
/FROM BDH /SOURCE VARIABLES=TIPHOGAR_m
/TARGET VARIABLES=TIPHOGAR_m
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
*----- comprobación -----.
RECODE TIPHOGAR_3 (SYSMIS=-9) (ELSE=Copy) INTO filas.

```

```

RECODE TIPHOGAR_m (SYSMIS=-9) (ELSE=Copy) INTO columnas.
execute.
APPLY DICTIONARY
  /FROM ArchivoH /SOURCE VARIABLES=TIPHOGAR_3
  /TARGET VARIABLES=filas
  /FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
APPLY DICTIONARY
  /FROM BDH /SOURCE VARIABLES=TIPHOGAR_m
  /TARGET VARIABLES=columnas
  /FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
CROSSTABS /TABLES=filas BY columnas /MISSING=INCLUDE.
DELETE VARIABLES filas columnas.

*=====,
* Principales datos resúmenes de Composicion del Hogar .
*=====,

* ----- n_dep: número de dependientes del hogar (según la ECV) -----,

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
AGGREGATE
  /OUTFILE=Aux
  /BREAK=idh
  /n_dep = SUM( dep ) .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
  /Table=Aux
  /BY idh .
execute.
APPLY DICTIONARY
  /FROM BDH /SOURCE VARIABLES=n_dep
  /TARGET VARIABLES=n_dep
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
*----- comprobación -----,
CROSSTABS n_dep by NMIEM7.

* ----- n_percep: número de perceptores del hogar -----,

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
AGGREGATE
  /OUTFILE=Aux
  /BREAK=idh
  /n_percep = SUM( perc ) .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
  /Table=Aux
  /BY idh .
execute.
APPLY DICTIONARY
  /FROM BDH /SOURCE VARIABLES=n_percep
  /TARGET VARIABLES=n_percep
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.

* ----- n_percep_: número de perceptores del hogar -----,

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
USE ALL.
COMPUTE filtro=(not missing(percep)).
FILTER BY filtro.
EXECUTE.
AGGREGATE
  /OUTFILE=Aux
  /BREAK=idh
  /N_casos = PIN ( percep 1 1 )
  /N_BREAK=N.
USE ALL.
FILTER OFF.
EXECUTE.
DELETE VARIABLES filtro .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
  /Table=Aux
  /BY idh .
execute.
compute N_casos= N_BREAK * N_casos / 100 .
compute n_percep_ = Rnd (N_casos) .

```

```

execute.
DELETE VARIABLES N_BREAK N_casos .
DATASET CLOSE Aux.
*----- comprobación -----.
CROSSTABS n_percep by n_percep_.
CROSSTABS n_percep by NUMPERI.

*----- n_hombre: número de hombres del hogar -----.

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
USE ALL.
COMPUTE filtro=(not missing(sexop)).
FILTER BY filtro.
EXECUTE.
AGGREGATE
  /OUTFILE=Aux
  /BREAK=idh
  /N_casos = PIN ( sexop 1 1 )
  /N_BREAK=N.
USE ALL.
FILTER OFF.
EXECUTE.
DELETE VARIABLES filtro .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
  /Table=Aux
  /BY idh .
execute.
compute N_casos= N_BREAK * N_casos / 100 .
compute n_hombre= Rnd (N_casos) .
execute.
APPLY DICTIONARY
  /FROM BDH /SOURCE VARIABLES=n_hombre
  /TARGET VARIABLES=n_hombre
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
DELETE VARIABLES N_BREAK N_casos .

```

```

*----- n_mujer: número de mujeres del hogar -----.

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
USE ALL.
COMPUTE filtro=(not missing(sexop)).
FILTER BY filtro.
EXECUTE.
AGGREGATE
  /OUTFILE=Aux
  /BREAK=idh
  /N_casos = PIN ( sexop 6 6 )
  /N_BREAK=N.
USE ALL.
FILTER OFF.
EXECUTE.
DELETE VARIABLES filtro .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
  /Table=Aux
  /BY idh .
execute.
compute N_casos= N_BREAK * N_casos / 100 .
compute n_mujer= Rnd (N_casos) .
execute.
APPLY DICTIONARY
  /FROM BDH /SOURCE VARIABLES=n_mujer
  /TARGET VARIABLES=n_mujer
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
DELETE VARIABLES N_BREAK N_casos .

```

```

*----- n_soltero: número de solteros del hogar -----.

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
USE ALL.
COMPUTE filtro=(not missing(estcivilp)).
FILTER BY filtro.
EXECUTE.
AGGREGATE
  /OUTFILE=Aux

```

```

/BREAK=idh
/N_casos = PIN ( estcivlp 1 1 )
/N_BREAK=N.
USE ALL.
FILTER OFF.
EXECUTE.
DELETE VARIABLES filtro .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
/Table=Aux
/BY idh .
execute.
compute N_casos= N_BREAK * N_casos / 100 .
compute n_soltero= Rnd (N_casos) .
execute.
APPLY DICTIONARY
/FROM BDH /SOURCE VARIABLES=n_soltero
/TARGET VARIABLES=n_soltero
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
DELETE VARIABLES N_BREAK N_casos .

```

* ----- n_casado: número de casados del hogar -----.

```

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
USE ALL.
COMPUTE filtro=(not missing(estcivlp)).
FILTER BY filtro.
EXECUTE.
AGGREGATE
/OUTFILE=Aux
/BREAK=idh
/N_casos = PIN ( estcivlp 2 2 )
/N_BREAK=N.
USE ALL.
FILTER OFF.
EXECUTE.
DELETE VARIABLES filtro .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
/Table=Aux
/BY idh .
execute.
compute N_casos= N_BREAK * N_casos / 100 .
compute n_casado= Rnd (N_casos) .
execute.
APPLY DICTIONARY
/FROM BDH /SOURCE VARIABLES=n_casado
/TARGET VARIABLES=n_casado
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
DELETE VARIABLES N_BREAK N_casos .

```

* ----- n_separado: número de separados del hogar -----.

```

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
USE ALL.
COMPUTE filtro=(not missing(estcivlp)).
FILTER BY filtro.
EXECUTE.
AGGREGATE
/OUTFILE=Aux
/BREAK=idh
/N_casos = PIN ( estcivlp 3 3 )
/N_BREAK=N.
USE ALL.
FILTER OFF.
EXECUTE.
DELETE VARIABLES filtro .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
/Table=Aux
/BY idh .
execute.
compute N_casos= N_BREAK * N_casos / 100 .
compute n_separado = Rnd (N_casos) .
execute.
APPLY DICTIONARY
/FROM BDH /SOURCE VARIABLES=n_separado

```

```

/TARGET VARIABLES=n_separado
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
DELETE VARIABLES N_BREAK N_casos .

```

* ----- n_viudo: número de viudos del hogar -----.

```

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
USE ALL.
COMPUTE filtro=(not missing(estcivlp)).
FILTER BY filtro.
EXECUTE.
AGGREGATE
/OUTFILE=Aux
/BREAK=idh
/N_casos = PIN ( estcivlp 4 4 )
/N_BREAK=N.
USE ALL.
FILTER OFF.
EXECUTE.
DELETE VARIABLES filtro .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
/Table=Aux
/BY idh .
execute.
compute N_casos= N_BREAK * N_casos / 100 .
compute n_viudo = Rnd (N_casos) .
execute.
APPLY DICTIONARY
/FROM BDH /SOURCE VARIABLES=n_viudo
/TARGET VARIABLES=n_viudo
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
DELETE VARIABLES N_BREAK N_casos .

```

* ----- n_divorciado: número de divorciados del hogar -----.

```

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
USE ALL.
COMPUTE filtro=(not missing(estcivlp)).
FILTER BY filtro.
EXECUTE.
AGGREGATE
/OUTFILE=Aux
/BREAK=idh
/N_casos = PIN ( estcivlp 5 5 )
/N_BREAK=N.
USE ALL.
FILTER OFF.
EXECUTE.
DELETE VARIABLES filtro .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
/Table=Aux
/BY idh .
execute.
compute N_casos= N_BREAK * N_casos / 100 .
compute n_divorciado= Rnd (N_casos) .
execute.
APPLY DICTIONARY
/FROM BDH /SOURCE VARIABLES=n_divorciado
/TARGET VARIABLES=n_divorciado
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
DELETE VARIABLES N_BREAK N_casos .

```

* ----- n_espanola: número de españoles del hogar -----.

```

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
USE ALL.
COMPUTE filtro=(not missing(nacip)).
FILTER BY filtro.
EXECUTE.
AGGREGATE
/OUTFILE=Aux
/BREAK=idh
/N_casos = PIN ( nacip 1 1 )

```

```

/N_BREAK=N.
USE ALL.
FILTER OFF.
EXECUTE.
DELETE VARIABLES filtro .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
  /Table=Aux
  /BY idh .
execute.
compute N_casos= N_BREAK * N_casos / 100 .
compute n_espanola= Rnd (N_casos) .
execute.
APPLY DICTIONARY
  /FROM BDH /SOURCE VARIABLES=n_espanola
  /TARGET VARIABLES=n_espanola
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
DELETE VARIABLES N_BREAK N_casos .

```

* ----- n_exUE24: número de miembros extranjeros del resto de la UE-25 del hogar -----.

```

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
USE ALL.
COMPUTE filtro=(not missing(nacip)).
FILTER BY filtro.
EXECUTE.
AGGREGATE
  /OUTFILE=Aux
  /BREAK=idh
  /N_casos = PIN ( nacip 2 2 )
  /N_BREAK=N.
USE ALL.
FILTER OFF.
EXECUTE.
DELETE VARIABLES filtro .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
  /Table=Aux
  /BY idh .
execute.
compute N_casos= N_BREAK * N_casos / 100 .
compute n_exUE24= Rnd (N_casos) .
execute.
APPLY DICTIONARY
  /FROM BDH /SOURCE VARIABLES=n_exUE24
  /TARGET VARIABLES=n_exUE24
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
DELETE VARIABLES N_BREAK N_casos .

```

* ----- n_exResE: número de miembros extranjeros del resto de Europa del hogar -----.

```

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
USE ALL.
COMPUTE filtro=(not missing(nacip)).
FILTER BY filtro.
EXECUTE.
AGGREGATE
  /OUTFILE=Aux
  /BREAK=idh
  /N_casos = PIN ( nacip 3 3 )
  /N_BREAK=N.
USE ALL.
FILTER OFF.
EXECUTE.
DELETE VARIABLES filtro .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
  /Table=Aux
  /BY idh .
execute.
compute N_casos= N_BREAK * N_casos / 100 .
compute n_exResE= Rnd (N_casos) .
execute.
APPLY DICTIONARY
  /FROM BDH /SOURCE VARIABLES=n_exResE
  /TARGET VARIABLES=n_exResE
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

```



```
DATASET CLOSE Aux.
DELETE VARIABLES N_BREAK N_casos .
```

* ----- n_exResM: número de extranjeros del resto del mundo en el hogar -----.

```
DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
USE ALL.
COMPUTE filtro=(not missing(nacip)).
FILTER BY filtro.
EXECUTE.
AGGREGATE
  /OUTFILE=Aux
  /BREAK=idh
  /N_casos = PIN ( nacip 4 4 )
  /N_BREAK=N.
USE ALL.
FILTER OFF.
EXECUTE.
DELETE VARIABLES filtro .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
  /Table=Aux
  /BY idh .
execute.
compute N_casos= N_BREAK * N_casos / 100 .
compute n_exResM= Rnd (N_casos) .
execute.
APPLY DICTIONARY
  /FROM BDH /SOURCE VARIABLES=n_exResM
  /TARGET VARIABLES=n_exResM
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
DELETE VARIABLES N_BREAK N_casos .
```

* ----- n_sinestud: número de miembros sin estudios del hogar -----.

```
DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
USE ALL.
COMPUTE filtro=(not missing(nedp)).
FILTER BY filtro.
EXECUTE.
AGGREGATE
  /OUTFILE=Aux
  /BREAK=idh
  /N_casos = PIN ( nedp 1 1 )
  /N_BREAK=N.
USE ALL.
FILTER OFF.
EXECUTE.
DELETE VARIABLES filtro .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
  /Table=Aux
  /BY idh .
execute.
compute N_casos= N_BREAK * N_casos / 100 .
compute n_sinestud= Rnd (N_casos) .
execute.
APPLY DICTIONARY
  /FROM BDH /SOURCE VARIABLES=n_sinestud
  /TARGET VARIABLES=n_sinestud
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
DELETE VARIABLES N_BREAK N_casos .
```

* ----- n_eduprimaria: número de miembros sin estudios o con Educación Primaria del hogar -----.

```
DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
USE ALL.
COMPUTE filtro=(not missing(nedp)).
FILTER BY filtro.
EXECUTE.
AGGREGATE
  /OUTFILE=Aux
  /BREAK=idh
  /N_casos = PIN ( nedp 2 2 )
  /N_BREAK=N.
USE ALL.
```

```

FILTER OFF.
EXECUTE.
DELETE VARIABLES filtro .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
  /Table=Aux
  /BY idh .
execute.
compute N_casos= N_BREAK * N_casos / 100 .
compute n_eduprimaria= Rnd (N_casos) .
execute.
APPLY DICTIONARY
  /FROM BDH /SOURCE VARIABLES=n_eduprimaria
  /TARGET VARIABLES=n_eduprimaria
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
DELETE VARIABLES N_BREAK N_casos .

```

* ----- n_edusec1: número de miembros con Educación Secundaria de 1ª etapa del hogar -----.

```

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
USE ALL.
COMPUTE filtro=(not missing(nedp)).
FILTER BY filtro.
EXECUTE.
AGGREGATE
  /OUTFILE=Aux
  /BREAK=idh
  /N_casos = PIN ( nedp 3 3 )
  /N_BREAK=N.
USE ALL.
FILTER OFF.
EXECUTE.
DELETE VARIABLES filtro .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
  /Table=Aux
  /BY idh .
execute.
compute N_casos= N_BREAK * N_casos / 100 .
compute n_edusec1= Rnd (N_casos) .
execute.
APPLY DICTIONARY
  /FROM BDH /SOURCE VARIABLES=n_edusec1
  /TARGET VARIABLES=n_edusec1
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
DELETE VARIABLES N_BREAK N_casos .

```

* ----- n_edusec2: número de miembros con Educación Secundaria de 2ª etapa del hogar -----.

```

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
USE ALL.
COMPUTE filtro=(not missing(nedp)).
FILTER BY filtro.
EXECUTE.
AGGREGATE
  /OUTFILE=Aux
  /BREAK=idh
  /N_casos = PIN ( nedp 4 4 )
  /N_BREAK=N.
USE ALL.
FILTER OFF.
EXECUTE.
DELETE VARIABLES filtro .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
  /Table=Aux
  /BY idh .
execute.
compute N_casos= N_BREAK * N_casos / 100 .
compute n_edusec2= Rnd (N_casos) .
execute.
APPLY DICTIONARY
  /FROM BDH /SOURCE VARIABLES=n_edusec2
  /TARGET VARIABLES=n_edusec2
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
DELETE VARIABLES N_BREAK N_casos .

```

* ----- n_edusuperior: número de miembros con Educación Superior del hogar -----.

```

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
USE ALL.
COMPUTE filtro=(not missing(nedp)).
FILTER BY filtro.
EXECUTE.
AGGREGATE
  /OUTFILE=Aux
  /BREAK=idh
  /N_casos = PIN ( nedp 5 5 )
  /N_BREAK=N.
USE ALL.
FILTER OFF.
EXECUTE.
DELETE VARIABLES filtro .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
  /Table=Aux
  /BY idh .
execute.
compute N_casos= N_BREAK * N_casos / 100 .
compute n_edusuperior= Rnd (N_casos) .
execute.
APPLY DICTIONARY
  /FROM BDH /SOURCE VARIABLES=n_edusuperior
  /TARGET VARIABLES=n_edusuperior
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
DELETE VARIABLES N_BREAK N_casos .

```

* ----- n_ocupado: número de ocupados del hogar -----.

```

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
USE ALL.
COMPUTE filtro=(not missing(situac3p)).
FILTER BY filtro.
EXECUTE.
AGGREGATE
  /OUTFILE=Aux
  /BREAK=idh
  /N_casos = PIN ( situac3p 1 1 )
  /N_BREAK=N.
USE ALL.
FILTER OFF.
EXECUTE.
DELETE VARIABLES filtro .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
  /Table=Aux
  /BY idh .
execute.
compute N_casos= N_BREAK * N_casos / 100 .
compute n_ocupado= Rnd (N_casos) .
execute.
APPLY DICTIONARY
  /FROM BDH /SOURCE VARIABLES=n_ocupado
  /TARGET VARIABLES=n_ocupado
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
DELETE VARIABLES N_BREAK N_casos .

```

* ----- n_ocup: número de ocupados del hogar -----.

```

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
AGGREGATE
  /OUTFILE=Aux
  /BREAK=idh
  /n_ocup_ = SUM( ocup ) .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
  /Table=Aux
  /BY idh .
execute.
DATASET CLOSE Aux.
*----- comprobacion ----.
CROSSTABS n_ocup_ BY n_ocupado.

```

* ----- n_desempleado: número de activos desempleados del hogar -----.

```

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
USE ALL.
COMPUTE filtro=(not missing(situac3p)).
FILTER BY filtro.
EXECUTE.
AGGREGATE
  /OUTFILE=Aux
  /BREAK=idh
  /N_casos = PIN ( situac3p 2 2 )
  /N_BREAK=N.
USE ALL.
FILTER OFF.
EXECUTE.
DELETE VARIABLES filtro .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
  /Table=Aux
  /BY idh .
execute.
compute N_casos= N_BREAK * N_casos / 100 .
compute n_desempleado= Rnd (N_casos) .
execute.
APPLY DICTIONARY
  /FROM BDH /SOURCE VARIABLES=n_desempleado
  /TARGET VARIABLES=n_desempleado
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
DELETE VARIABLES N_BREAK N_casos .

```

* ----- n_estudiante: número de inactivos estudiantes del hogar -----.

```

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
USE ALL.
COMPUTE filtro=(not missing(situac3p)).
FILTER BY filtro.
EXECUTE.
AGGREGATE
  /OUTFILE=Aux
  /BREAK=idh
  /N_casos = PIN ( situac3p 3 3 )
  /N_BREAK=N.
USE ALL.
FILTER OFF.
EXECUTE.
DELETE VARIABLES filtro .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
  /Table=Aux
  /BY idh .
execute.
compute N_casos= N_BREAK * N_casos / 100 .
compute n_estudiante= Rnd (N_casos) .
execute.
APPLY DICTIONARY
  /FROM BDH /SOURCE VARIABLES=n_estudiante
  /TARGET VARIABLES=n_estudiante
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
DELETE VARIABLES N_BREAK N_casos .

```

* ----- n_retirado: número de inactivos jubilados o retirados del hogar -----.

```

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
USE ALL.
COMPUTE filtro=(not missing(situac3p)).
FILTER BY filtro.
EXECUTE.
AGGREGATE
  /OUTFILE=Aux
  /BREAK=idh
  /N_casos = PIN ( situac3p 4 4 )
  /N_BREAK=N.
USE ALL.
FILTER OFF.
EXECUTE.

```

```

DELETE VARIABLES filtro .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
  /Table=Aux
  /BY idh .
execute.
compute N_casos= N_BREAK * N_casos / 100 .
compute n_retirado= Rnd (N_casos) .
execute.
APPLY DICTIONARY
  /FROM BDH /SOURCE VARIABLES=n_retirado
  /TARGET VARIABLES=n_retirado
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
DELETE VARIABLES N_BREAK N_casos .

```

* ----- n_labhogar: número inactivos dedicados a labores del hogar -----.

```

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
USE ALL.
COMPUTE filtro=(not missing(situac3p)).
FILTER BY filtro.
EXECUTE.
AGGREGATE
  /OUTFILE=Aux
  /BREAK=idh
  /N_casos = PIN ( situac3p 5 5 )
  /N_BREAK=N.
USE ALL.
FILTER OFF.
EXECUTE.
DELETE VARIABLES filtro .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
  /Table=Aux
  /BY idh .
execute.
compute N_casos= N_BREAK * N_casos / 100 .
compute n_labhogar= Rnd (N_casos) .
execute.
APPLY DICTIONARY
  /FROM BDH /SOURCE VARIABLES=n_labhogar
  /TARGET VARIABLES=n_labhogar
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
DELETE VARIABLES N_BREAK N_casos .

```

* ----- n_otrinactivo: número de inactivos del hogar -----.

```

DATASET DECLARE Aux WINDOW=HIDDEN.
DATASET ACTIVATE ArchivoP.
USE ALL.
COMPUTE filtro=(not missing(situac3p)).
FILTER BY filtro.
EXECUTE.
AGGREGATE
  /OUTFILE=Aux
  /BREAK=idh
  /N_casos = PIN ( situac3p 6 6 )
  /N_BREAK=N.
USE ALL.
FILTER OFF.
EXECUTE.
DELETE VARIABLES filtro .
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
  /Table=Aux
  /BY idh .
execute.
compute N_casos= N_BREAK * N_casos / 100 .
compute n_otrinactivo= Rnd (N_casos) .
execute.
APPLY DICTIONARY
  /FROM BDH /SOURCE VARIABLES=n_otrinactivo
  /TARGET VARIABLES=n_otrinactivo
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
DATASET CLOSE Aux.
DELETE VARIABLES N_BREAK N_casos .

```

Value labels

```
nmh nmh16 nmh14 numhabit_m n_dep n_percep n_hombre n_mujer n_soltero n_casado n_viudo n_separado n_divorciado
n_espanola n_exue24 n_exrese n_exresM n_sinstud n_eduprimaria n_edusec1 n_edusec2 n_edusuperior
n_ocupado n_desempleado n_estudiante n_retirado n_labhogar n_otrinactivo
0'0' 1'1' 2'2' 3'3' 4'4' 5'5' 6'6' 7'7' 8'8' 9'9' 10'10' 11'11' 12'12' 13'13' 14'14' 15'15' 16'16' 17'17' 18'18' 19'19'.
```

```
*-----.
```

```
ECHO ">>> comprobaciones de variables resúmenes de hogares".
```

```
DATASET ACTIVATE ArchivoH.
compute diff_sexo = nmh-(n_hombre+n_mujer).
compute diff_estcivil = nmh16-(n_soltero+n_casado+n_viudo+n_separado+n_divorciado).
compute diff_naci = nmh16-(n_espanola+n_exUE24+n_exResE +n_exResM).
compute diff_nestud = nmh16-(n_sinstud+n_eduprimaria+n_edusec1+n_edusec2+n_edusuperior).
compute diff_situac = nmh16-(n_ocupado+n_desempleado+n_estudiante+n_retirado+n_labhogar+n_otrinactivo) .
execute.
FRECUENCIES VARIABLES=diff_sexo diff_estcivil diff_naci diff_nestud diff_situac .
FRECUENCIES VARIABLES=n_hombre n_mujer n_soltero n_casado n_viudo n_separado n_divorciado n_espanola n_exUE24
n_exResE n_exResM
n_sinstud n_eduprimaria n_edusec1 n_edusec2 n_edusuperior n_ocupado n_desempleado n_estudiante n_retirado
n_labhogar n_otrinactivo
/ORDER=ANALYSIS.
```

```
DELETE VARIABLES diff_sexo diff_estcivil diff_naci diff_nestud diff_situac.
```

```
*=====.
```

```
** Variables sobre los Ingresos del Hogar .
```

```
*=====.
```

```
*----- MATCH - INTOTALES_m: Total de Ingresos Netos del Hogar -----.
```

```
DATASET ACTIVATE ArchivoH .
if not(missing(IMPEXAC)) INTOTALES_m=SUM(IMPEXAC*12, GASTNOM1/FACTOR, GASTNOM2/FACTOR, GASTNOM3/FACTOR).
EXECUTE.
APPLY DICTIONARY
/FROM BDH /SOURCE VARIABLES=INTOTALES_m
/TARGET VARIABLES=INTOTALES_m
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
```

```
*----- MATCH - Pi_INTOTALES_m y IIC10_INTOTALES_m: Orden de cuantil (pi) y Grupo intercuantilico (a 10 IIC)
```

```
* según el el total de ingresos netos del Hogar -----.
```

```
!Calcula_Grupos_Intercuanticlicos segm=ano peso=w var=INTOTALES_m ngrupos=10 .
```

```
*----- MATCH - LINTOTALES_m: Ingresos Netos del Hogar -----.
```

```
DATASET ACTIVATE ArchivoH .
compute LINTOTALES_m=lg10( (INTOTALES_m<=0)+INTOTALES_m*(INTOTALES_m>0) ).
execute.
APPLY DICTIONARY
/FROM BDH /SOURCE VARIABLES=LINTOTALES_m
/TARGET VARIABLES=LINTOTALES_m
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
```

```
*----- MATCH - IIC10_LINTOTALES_m: Grupo intercuantilico (a 10 IIC)
```

```
* según el el total del log de los ingresos netos del Hogar -----.
```

```
!Calcula_Grupos_Intercuanticlicos segm=ano peso=w var=lintotales_m ngrupos=10 pi=no .
```

```
*----- MATCH - ININTOTALES_m: Intervalos de Ingresos Netos del Hogar -----.
```

```
DATASET ACTIVATE ArchivoH .
RECODE
INTOTALES_m
(MISSING=SYSMIS) (LOWEST thru -0.001=0) (0 thru 5999.999=1) (6000 thru 11999.999=2) (12000 thru 17999.999=3) (18000 thru
23999.999=4)
(24000 thru 29999.999=5) (30000 thru 35999.999=6) (36000 thru 59999.999=7) (60000 thru 83999.999=8) (84000 thru
107999.999=9)
(108000 thru HIGHEST =10) INTO ININTOTALES_m.
execute.
APPLY DICTIONARY
/FROM BDH /SOURCE VARIABLES=ININTOTALES_m
/TARGET VARIABLES=ININTOTALES_m
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
```

```
*----- MATCH - IND1_IT_m: Índice 1 del Total de Ingresos Netos del Hogar -----.
```

```
DATASET DECLARE Aux WINDOW=HIDDEN.
```

```
DATASET ACTIVATE ArchivoH.
```

```
USE ALL.
```

```
COMPUTE lng_N = INTOTALES_m.
```

```
EXECUTE.
```

```
COMPUTE filtro=(not missing(lng_N)).
```

```
FILTER BY filtro.
```

```

EXECUTE.
AGGREGATE
/OUTFILE=Aux
/BREAK= pais ano
/ing_N_sd = SD(Ing_N)
/ing_N_median = MEDIAN(Ing_N) .
USE ALL.
FILTER OFF.
EXECUTE.
DATASET ACTIVATE ArchivoH .
MATCH FILES /FILE=*
/Table=Aux
/BY pais ano .
execute.
compute      IND0_IT_m      =      (Ing_N-Ing_N_median)      /      (      ((3*Ing_N_sd-
Ing_N_median)/(3*Ing_N_sd+Ing_N_median))*max(0,min(Ing_N,Ing_N_median+3*Ing_N_sd)) +Ing_N_median ).
compute IND1_IT_m = max(-1,IND0_IT_m).
execute.
formats IND0_IT_m IND1_IT_m (F10,3).
VARIABLE LABELS  IND0_IT_m 'Indice 0 del Total de Ingresos Netos del Hogar'
IND1_IT_m 'Indice 1 del Total de Ingresos Netos del Hogar'.
DATASET CLOSE Aux.
DELETE VARIABLES filtro ing_N ing_N_median ing_N_sd .

*----- MATCH - IIC10_IND1_IT_m: Grupo intercuantilico (a 10 IIC)
*      según el IND1_IT_m de los ingresos netos del Hogar -----,
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=IND1_IT_m ngrupos=10 pi=no .

*----- Gasto Total anual del Hogar -----,

COMPUTE GASTOH=GASTOT/FACTOR.
VARIABLE LABELS  GASTOH 'Gasto Total anual del hogar'.
EXECUTE.

*----- MATCH - Pi_GASTOH y IIC10_GASTOH: Orden de cuantil (pi) y Grupo intercuantilico (a 10 IIC)
*      según el Gasto Total del Hogar -----,
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=GASTOH ngrupos=10 .

*----- Índice de Cobertura -----,

if (INTOTALES_m>0) IC=GASTOH/INTOTALES_m.
VARIABLE LABELS  IC 'Índice de cobertura '.
EXECUTE.

*----- MATCH - Pi_IC y IIC10_IC: Orden de cuantil (pi) y Grupo intercuantilico (a 10 IIC)
*      según el Índice de Cobertura -----,
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=IC ngrupos=10 .

* ---- MATCH - Fuentes de ingresos del Hogar -----,

DATASET ACTIVATE ArchivoH .
RECODE CAPROP (-9=SYSMIS) (MISSING=SYSMIS) (1=1) (6=0) INTO I_CPRO.
RECODE CAJENA (-9=SYSMIS) (MISSING=SYSMIS) (1=1) (6=0) INTO I_CAJE.
RECODE PENSIO (-9=SYSMIS) (MISSING=SYSMIS) (1=1) (6=0) INTO I_PENS.
RECODE DESEM (-9=SYSMIS) (MISSING=SYSMIS) (1=1) (6=0) INTO I_DES.
RECODE OTRSUB (-9=SYSMIS) (MISSING=SYSMIS) (1=1) (6=0) INTO I_OSPS.
RECODE RENTAS (-9=SYSMIS) (MISSING=SYSMIS) (1=1) (6=0) INTO I_RCYP.
RECODE OTROIN (-9=SYSMIS) (MISSING=SYSMIS) (1=1) (6=0) INTO I_OTRI.
EXECUTE.
VARIABLE LABELS
I_DES  'El hogar tiene Ingresos por Desempleo'
I_OSPS 'El hogar tiene Ingresos por Otras Subvenciones y PS'
I_RCYP 'El hogar tiene Ingresos por Rentas del Capital y la Propiedad'
I_OTRI 'El hogar tiene Ingresos por Otros Ingresos'
I_CAJE 'El hogar tiene Ingresos por Cuenta Ajena'
I_CPRO 'El hogar tiene Ingresos por Cuenta Propia'
I_PENS 'El hogar tiene Ingresos por Pensiones C y NC'.
VALUE LABELS I_CPRO I_CAJE I_PENS I_DES I_OSPS I_RCYP I_OTRI
1 'Sí'
0 'No'.

* ---- MATCH -FP_ING_m: Fuente Principal de ingresos del Hogar -----,

DATASET ACTIVATE ArchivoH .
RECODE FUENPRIN (-9 MISSING=SYSMIS) (ELSE=Copy) INTO FP_ing_m.
execute.
APPLY DICTIONARY
/FROM BDH /SOURCE VARIABLES=FP_ing_m
/TARGET VARIABLES=FP_ing_m
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

```

```

RECODE FP_ing_m (-9 MISSING=SYSMIS) (4 5 6 7 = 4) (ELSE=Copy) INTO FP_ing2_m.
execute.
Formats FP_ING2_m (F2.0).
VARIABLE LABELS FP_ING2_m 'Fuente Principal de los ingresos regulares en el hogar'.
VALUE LABELS FP_ING2_m
1 'Cuenta Propia'
2 'Cuenta Ajena'
3 'Pensiones Contributivas y no contributivas'
4 'Otros ingresos'.

```

```

*=====,
* Identificación de las personas mas características del Hogar .
*=====,

```

```

DATASET ACTIVATE ArchivoP .
SORT CASES BY idh (A) norden_ing (A).
execute.

```

```

* ----- Identificación del Sustentador Principal (SP) -----,

```

```

DATASET ACTIVATE ArchivoH .
compute norden_ing=1.
execute.
MATCH FILES /FILE=*
/Table='ArchivoP'
/RENAME idp = cod_SP
/BY idh norden_ing
/KEEP= ANOENC to FP_ing2_m cod_SP .
execute.
APPLY DICTIONARY
/FROM BDH /SOURCE VARIABLES=cod_SP
/TARGET VARIABLES=cod_SP
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

```

***** Ojo, el SP en las EPF no es el que más aporta al hogar. Posible mejora, adoptando la definición de SP de la EPF en la ECV.

```

DATASET ACTIVATE ArchivoH .
COMPUTE cod_SP_=100*NUMERO+NORDENSP.
EXECUTE.

```

```

compute diff_Cod_SP= not (cod_SP = cod_SP_).
execute.

```

```

* ----- Identificación del segundo Adulto (2A) -----,

```

*** (El 2A es el que más ingresos aporta al Hogar, después del SP; y en caso de empate, el de mayor edad (con preferencia a los activos cuando la aportación es negativa o nula)

```

DATASET ACTIVATE ArchivoH .
compute norden_ing=2.
execute.
MATCH FILES /FILE=*
/Table='ArchivoP'
/RENAME idp = cod_2A
/BY idh norden_ing
/KEEP= ANOENC to diff_Cod_SP cod_2A .
execute.
Formats cod_2A (F10.0).
Variable labels cod_2A 'Número de identificación del 2º Adulto del Hogar'.

```

```

DATASET ACTIVATE ArchivoP .
SORT CASES BY idp (A) .
execute.

```

```

DATASET ACTIVATE ArchivoH .
compute diff_Cod_2A= (cod_2A = cod_SP_).
execute.

```

```

*=====,
** Carga de datos del Sustentador Principal (SP) desde el archivo de personas.
*=====,

```

```

DATASET ACTIVATE ArchivoH .
SORT CASES BY cod_SP (A) .
execute.
MATCH FILES /FILE=*
/Table='ArchivoP'
/RENAME idp = cod_SP
sexop = sexo_SP
edadp = edad_SP

```



```

gedadp = edad_SP
nedp = ned_SP
nacip = naci_SP
estcivilp = estcivil_SP
csep = cse_SP
ocupacp = ocupac_SP
activestb1p = activestb1_SP
activestb2p = activestb2_SP
situprof = situprof_SP
tipcontp = tipcont_SP
situac3p = situac3_SP
/BY cod_SP
/KEEP= ANOENC to diff_Cod_2A sexo_SP edad_SP edad_SP estcivil_SP ned_SP
      naci_SP cse_SP ocupac_SP activestb1_SP activestb2_SP situprof_SP tipcont_SP situac3_SP .
execute.

SORT CASES BY idh (A) .
execute.
=====
** Datos del Sustentador Principal (SP) tomados del archivo de Hogares .
=====

DATASET ACTIVATE ArchivoH .
STRING VAux (A5).

DO IF (cod_SP = cod_SP_).

    RECODE SITUACT_SP (SYSMIS=-2) (-9=SYSMIS) (1 2=1) (3=2) (4=4) (5=3) (6=5) (7=6) (ELSE=Copy) INTO situac3_SP.

    RECODE TIPOCONT (SYSMIS=-2) (-9=SYSMIS) (2 3=2) (ELSE=Copy) INTO tipcont_SP.
    if ((SITPROF ge 2 and SITPROF le 4) or (SITUACT_SP ge 3 and SITUACT_SP le 7)) tipcont_SP=-2 /* definido el la ECV solo para
    ocupados-asalariados */.
    if (gedad_SP=0) tipcont_SP=0 /* 0 ND (menor de 16 años) */.

    RECODE SITPROF (SYSMIS=-2) (-9=SYSMIS) (3=1) (2=2) (1=3) (4=4) (ELSE=Copy) INTO situprof_SP.
    if (SITUACT_SP ge 3 and SITUACT_SP le 7) situprof_SP=-2 /* definido el la ECV solo para ocupados */.
    if (gedad_SP=0) situprof_SP=0 /* 0 ND (menor de 16 años) */.

    RECODE OCUPA (SYSMIS=-2) (-9=SYSMIS) (0=10) (ELSE=Copy) INTO ocupac_SP.
    if (SITUACT_SP ge 3 and SITUACT_SP le 7) ocupac_SP=-2 /* definido el la ECV solo para ocupados */.
    if (gedad_SP=0) ocupac_SP=0 /* 0 ND (menor de 16 años) */.

    compute VAux=LTRIM(RTRIM(ATESTB)).
    if (SITUACT_SP ge 3 and SITUACT_SP le 7) VAux=' ' /* definido el la ECV solo para ocupados */.

lif (!lepf !le 'EPF08') !then
    RECODE VAux (' '= '_NA') ('-9'=' ') ('A' 'B'='A+B') ('C' 'D' 'E'='C+D+E') ('O' 'P' 'Q'='O+P+Q')
    (ELSE=Copy) INTO activestb1_SP.
lifend

lif (!lepf !ge 'EPF09') !then
    RECODE VAux (' '= '_NA') ('-9'=' ') (ELSE=Copy) INTO activestb2_SP.
    RECODE activestb2_SP ('A'='A+B') ('B' 'C' 'D' 'E'='C+D+E') ('F'='F') ('G'='G') ('I'='H') ('H' 'J'='I')
    ('K'='J') ('L'='K') ('O'='L') ('P'='M') ('Q'='N') ('M' 'N' 'R' 'S' 'T' 'U'='O+P+Q')
    (ELSE=Copy) INTO activestb1_SP.
lifend

compute cse_SP=$SYSMIS.
if ((situac3_SP = 1) and (situprof_SP = 1) and (ATESTB eq 'A')) cse_SP=1 .
if ((situac3_SP = 1) and (situprof_SP = 1) and (ATESTB ne " and ATESTB ne '-9' and ATESTB ne 'A')) cse_SP=2 .
if ((situac3_SP = 1) and (situprof_SP = 1) and (ATESTB eq " or ATESTB eq '-9')) cse_SP=3 .
if ((situac3_SP = 1) and (situprof_SP = 2) and (ATESTB = 'A')) cse_SP=4 .
if ((situac3_SP = 1) and (situprof_SP = 2) and (ATESTB ne " and ATESTB ne '-9' and ATESTB ne 'A')) cse_SP=5 .
if ((situac3_SP = 1) and (situprof_SP = 2) and (ATESTB eq " or ATESTB eq '-9')) cse_SP=6 .
if ((situac3_SP = 1) and (3 <= situprof_SP and situprof_SP <= 4) and (ocupac_SP eq 6 or ATESTB = 'A')) cse_SP=7 .
if ((situac3_SP = 1) and (3 <= situprof_SP and situprof_SP <= 4) and (1<=ocupac_SP and ocupac_SP<=2) and (ATESTB ne "
and ATESTB ne '-9' and ATESTB ne 'A')) cse_SP=8 .
if ((situac3_SP = 1) and (3 <= situprof_SP and situprof_SP <= 4) and (3<=ocupac_SP and ocupac_SP<=8) and (ocupac_SP
ne 6) and (ATESTB ne " and ATESTB ne '-9' and ATESTB ne 'A')) cse_SP=9 .
if ((situac3_SP = 1) and (3 <= situprof_SP and situprof_SP <= 4) and (ocupac_SP=9) and (ATESTB ne " and ATESTB ne '-9'
and ATESTB ne 'A')) cse_SP=10 .
if ((situac3_SP = 1) and (3 <= situprof_SP and situprof_SP <= 4) and (ocupac_SP=10) and (ATESTB ne " and ATESTB ne '-9'
and ATESTB ne 'A')) cse_SP=11 .
if ((situac3_SP = 1) and (3 <= situprof_SP and situprof_SP <= 4) and (ATESTB eq " or ATESTB eq '-9')) cse_SP=12 .
if (situac3_SP = 2) cse_SP=13 .
if (situac3_SP = 3) cse_SP=14 .
if (situac3_SP = 4) cse_SP=15 .
if (situac3_SP = 5 or situac3_SP = 6) cse_SP=16 .
if (gedad_SP=0) cse_SP=0 .

END IF.

```

```

execute.
DELETE VARIABLES VAux.

*-----
ECHO ">>> comprobaciones Sustentador Principal".

DATASET ACTIVATE ArchivoH .
FRECUENCIAS VARIABLES=sexo_SP edad_SP gedad_SP ned_SP naci_SP estcivil_SP
cse_SP ocupac_SP activestb1_SP activestb2_SP situatprof_SP tipcont_SP situac3_SP /ORDER=ANALYSIS.

CROSSTABS sexo_SP by SEXOSP.
CROSSTABS edad_SP by EDADSP.
CROSSTABS ned_SP by ESTUDRED_SP.
CROSSTABS ned_SP by ESTUDIOS_SP.
CROSSTABS naci_SP by NACIONAL_SP.
!if (!pf !le 'EPF10' ) !then
CROSSTABS naci_SP by PAISSP.
CROSSTABS estcivil_SP by ECIVILSP.
!lifend
!if (!pf !ge 'EPF11' ) !then
CROSSTABS naci_SP by PAISCOD_SP.
CROSSTABS estcivil_SP by ECIVILLEGALSP.
!lifend
CROSSTABS situac3_SP by SITUACT_SP.

*----- otras variables del SP:
* PERCEPSP IMPEXACP_SP IMPUEXACP_SP INTERINP_SP IMPU_INTERPSP.
* TRABAJO OCUPA OCUPARED SECTOR SITSOCI SITSOCI_RE.

!if (!pf !ge 'EPF09' ) !then
*---- activestb2_SP_m: Actividad del establecimiento laboral (2ª clasif) -----
*----- (a partir del 2010) Recodificación numérica -----
RECODE activestb2_SP (""='SYSMIS') (_NA'=-2) (_ND'=0) ('A'=1) ('B'=2) ('C'=3) ('D'=4) ('E'=5) ('F'=6) ('G'=7)
('H'=8) ('I'=9) ('J'=10) ('K'=11) ('L'=12) ('M'=13) ('N'=14) ('O'=15) ('P'=16)
('Q'=17) ('R'=18) ('S'=19) ('T'=20) ('U'=21) INTO activestb2_SP_m.
if (SITUACT_SP ge 3 and SITUACT_SP le 7) activestb2_SP_m =-2 /* definido el la ECV solo para ocupados */.
if (gedad_SP=0) activestb2_SP_m =0 /* 0 ND (menor de 16 años) */.
EXECUTE.
Variable level activestb2_SP_m (Nominal).
Variable labels activestb2_SP_m 'Actividad del establecimiento del que depende o dependía laboralmente'.
Value labels activestb2_SP_m
-2 'No Aplicable (Parado o Inactivo)'
0 'ND (menor de 16 años)'
1 'Agricultura, ganadería, silvicultura y pesca'
2 'Industrias extractivas'
3 'Industria manufacturera'
4 'Suministro de energía eléctrica, gas, vapor y aire acondicionado'
5 'Suministro de agua, actividades de saneamiento, gestión de residuos y descontaminación'
6 'Construcción'
7 'Comercio al por mayor y al por menor; reparación de vehículos de motor y motocicletas'
8 'Transporte y almacenamiento'
9 'Hostelería'
10 'Información y comunicaciones'
11 'Actividades financieras y de seguros'
12 'Actividades inmobiliarias'
13 'Actividades profesionales, científicas y técnicas'
14 'Actividades administrativas y servicios auxiliares'
15 'Administración pública y defensa; seguridad social obligatoria'
16 'Educación'
17 'Actividades sanitarias y de servicios sociales'
18 'Actividades artísticas, recreativas y de entretenimiento'
19 'Otros servicios'
20 'Actividades de los hogares como empleadores de personal doméstico'
21 'Actividades de organizaciones y organismos extraterritoriales'.
!lifend

*---- activestb1_SP_m: Actividad del establecimiento laboral (1ª clasif) -----
*----- (hasta el año 2009) Recodificación numérica -----
DATASET ACTIVATE ArchivoH .
RECODE activestb1_SP (""='SYSMIS') (_NA'=-2) (_ND'=0) ('A+B'=1) ('C+D+E'=2) ('F'=3) ('G'=4) ('H'=5) ('I'=6) ('J'=7)
('K'=8) ('L'=9) ('M'=10) ('N'=11) ('O+P+Q'=12) INTO activestb1_SP_m.
if (SITUACT_SP ge 3 and SITUACT_SP le 7) activestb1_SP_m =-2 /* definido el la ECV solo para ocupados */.
if (gedad_SP=0) activestb1_SP_m =0 /* 0 ND (menor de 16 años) */.
EXECUTE.
Variable level activestb1_SP_m (Nominal).
Variable labels activestb1_SP_m 'Actividad del establecimiento del que depende o dependía laboralmente'.
Value labels activestb1_SP_m
-2 'No Aplicable (Parado o Inactivo)'
0 'ND (menor de 16 años)'
1 'Agricultura, ganadería, caza, silvicultura y pesca'
2 'Indus. extractivas y manufact. Prod. y Distrib. de energía eléctrica, gas y agua'

```

3 'Construcción'
 4 'Comercio; reparación de vehículos de motor, ciclomotores y artículos de uso doméstico'
 5 'Hostelería'
 6 'Transporte, almacenamiento y comunicaciones'
 7 'Intermediación financiera'
 8 'Actividades inmobiliarias y de alquiler; Servicios empresariales'
 9 'Administración pública y defensa; seguridad social obligatoria'
 10 'Educación'
 11 'Actividades sanitarias y veterinarias; servicios sociales'
 12 'Otras Activ y de servicios a la comunidad ; Serv personales. Hogares que emplea p.doméstico. Organ extraterritoriales'.

```
=====,
** Datos del 2º Adulto en importancia de aportacion al Hogar (2A) tomados del archivo de personas .
=====,
```

```
DATASET ACTIVATE ArchivoP .
SORT CASES BY idp (A) .
execute.
```

```
DATASET ACTIVATE ArchivoH .
SORT CASES BY cod_2A (A) .
execute.
MATCH FILES /FILE=*
/Table='ArchivoP'
/RENAME idp = cod_2A
sexop = sexo_2A
edadp = edad_2A
gedadp = gedad_2A
nacip = naci_2A
nedp = ned_2A
estcivilp = estcivil_2A
csep = cse_2A
ocupacp = ocupac_2A
activestb1p = activestb1_2A
activestb2p = activestb2_2A
situprof = situprof_2A
tipcontp = tipcont_2A
situac3p = situac3_2A
/BY cod_2A
/KEEP= ANOENC to activestb1_SP_m sexo_2A edad_2A gedad_2A naci_2A ned_2A estcivil_2A
cse_2A ocupac_2A activestb1_2A activestb2_2A situprof_2A tipcont_2A situac3_2A .
execute.
```

```
SORT CASES BY idh (A) .
execute.
```

```
=====,
** Datos del 2º Adulto en importancia de aportacion al Hogar (2A) del archivo de Hogares .
=====,
```

```
DATASET ACTIVATE ArchivoH .
STRING VAux (A5).
```

```
DO IF (cod_2A = cod_SP_).
```

```
RECODE SITUACT_SP (SYSMIS=-2) (-9=SYSMIS) (1 2=1) (3=2) (4=4) (5=3) (6=5) (7=6) (ELSE=Copy) INTO situac3_2A.
```

```
RECODE TIPOCONT (SYSMIS=-2) (-9=SYSMIS) (2 3=2) (ELSE=Copy) INTO tipcont_2A.
if ((SITPROF ge 2 and SITPROF ge 4) or (SITUACT_SP ge 3 and SITUACT_SP le 7)) tipcont_2A=-2 /* definido el la ECV solo
para ocupados-asalariados */.
if (gedad_2A=0) tipcont_2A=0 /* 0 ND (menor de 16 años) */.
```

```
RECODE SITPROF (SYSMIS=-2) (-9=SYSMIS) (3=1) (2=2) (1=3) (4=4) (ELSE=Copy) INTO situprof_2A.
if (SITUACT_SP ge 3 and SITUACT_SP le 7) situprof_2A=-2 /* definido el la ECV solo para ocupados */.
if (gedad_2A=0) situprof_2A=0 /* 0 ND (menor de 16 años) */.
```

```
RECODE OCUPA (SYSMIS=-2) (-9=SYSMIS) (0=10) (ELSE=Copy) INTO ocupac_2A.
if (SITUACT_SP ge 3 and SITUACT_SP le 7) ocupac_2A=-2 /* definido el la ECV solo para ocupados */.
if (gedad_2A=0) ocupac_2A=0 /* 0 ND (menor de 16 años) */.
```

```
compute VAux=LTRIM(RTRIM(ATESTB)).
if (SITUACT_SP ge 3 and SITUACT_SP le 7) VAux=' ' /* definido el la ECV solo para ocupados */.
```

```
lif (!lef !le 'EPF08') !then
RECODE VAux ('='_NA') ('-9'=' ') ('A' 'B'='A+B') ('C' 'D' 'E'='C+D+E') ('O' 'P' 'Q'='O+P+Q')
(ELSE=Copy) INTO activestb1_2A.
```

```
lifend
```

```
lif (!lef !ge 'EPF09') !then
RECODE VAux ('='_NA') ('-9'=' ') (ELSE=Copy) INTO activestb2_2A.
```

```

RECODE activestb2_2A ('A'='A+B') ('B' 'C' 'D' 'E'='C+D+E') ('F'='F') ('G'='G') ('I'='H') ('H' 'J'='I')
('K'='J') ('L'='K') ('O'='L') ('P'='M') ('Q'='N') ('M' 'N' 'R' 'S' 'T' 'U'='O+P+Q')
(ELSE=Copy) INTO activestb1_2A.

lifend

compute cse_2A=$SYSMIS.
if ((situac3_2A = 1) and (situprof_2A = 1) and (ACTESTB eq 'A')) cse_2A=1 .
if ((situac3_2A = 1) and (situprof_2A = 1) and (ACTESTB ne " and ACTESTB ne '-9' and ACTESTB ne 'A' )) cse_2A=2 .
if ((situac3_2A = 1) and (situprof_2A = 1) and (ACTESTB eq " or ACTESTB eq '-9')) cse_2A=3 .
if ((situac3_2A = 1) and (situprof_2A = 2) and (ACTESTB = 'A')) cse_2A=4 .
if ((situac3_2A = 1) and (situprof_2A = 2) and (ACTESTB ne " and ACTESTB ne '-9' and ACTESTB ne 'A' )) cse_2A=5 .
if ((situac3_2A = 1) and (situprof_2A = 2) and (ACTESTB eq " or ACTESTB eq '-9')) cse_2A=6 .
if ((situac3_2A = 1) and (3 <= situprof_2A and situprof_2A <= 4) and (ocupac_2A eq 6 or ACTESTB = 'A' )) cse_2A=7 .
if ((situac3_2A = 1) and (3 <= situprof_2A and situprof_2A <= 4) and (1 <= ocupac_2A <= 2) and (ACTESTB ne " and ACTESTB ne
'-9' and ACTESTB ne 'A' )) cse_2A=8 .
if ((situac3_2A = 1) and (3 <= situprof_2A and situprof_2A <= 4) and (3 <= ocupac_2A <= 8) and (ocupac_2A ne 6) and
(ACTESTB ne " and ACTESTB ne '-9' and ACTESTB ne 'A' )) cse_2A=9 .
if ((situac3_2A = 1) and (3 <= situprof_2A and situprof_2A <= 4) and (ocupac_2A=9) and (ACTESTB ne " and ACTESTB ne '-9'
and ACTESTB ne 'A' )) cse_2A=10 .
if ((situac3_2A = 1) and (3 <= situprof_2A and situprof_2A <= 4) and (ocupac_2A=10) and (ACTESTB ne " and ACTESTB ne '-9'
and ACTESTB ne 'A' )) cse_2A=11 .
if ((situac3_2A = 1) and (3 <= situprof_2A and situprof_2A <= 4) and (ACTESTB eq " or ACTESTB eq '-9')) cse_2A=12 .
if (situac3_2A = 2) cse_2A=13 .
if (situac3_2A = 3) cse_2A=14 .
if (situac3_2A = 4) cse_2A=15 .
if (situac3_2A = 5 or situac3_2A = 6) cse_2A=16 .
if (gedad_2A=0) cse_2A=0 .

END IF.
execute.
DELETE VARIABLES VAux.

*=====,
** Etiquetado de los Datos del 2º Adulto en importancia de aportacion al Hogar (2A) .
*=====,

APPLY DICTIONARY
/ FROM BDH /SOURCE VARIABLES=sexo_SP
/TARGET VARIABLES=sexo_2A
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

APPLY DICTIONARY
/ FROM BDH /SOURCE VARIABLES=edad_SP
/TARGET VARIABLES=edad_2A
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

APPLY DICTIONARY
/ FROM BDH /SOURCE VARIABLES=gedad_SP
/TARGET VARIABLES=gedad_2A
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

APPLY DICTIONARY
/ FROM BDH /SOURCE VARIABLES=naci_SP
/TARGET VARIABLES=naci_2A
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

APPLY DICTIONARY
/ FROM BDH /SOURCE VARIABLES=ned_SP
/TARGET VARIABLES=ned_2A
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

APPLY DICTIONARY
/ FROM BDH /SOURCE VARIABLES=estcivil_SP
/TARGET VARIABLES=estcivil_2A
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

APPLY DICTIONARY
/ FROM BDH /SOURCE VARIABLES=activestb1_SP
/TARGET VARIABLES=activestb1_2A
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

lif (lepf !ge 'EPF09') !then
  APPLY DICTIONARY
  / FROM BDH /SOURCE VARIABLES=activestb2_SP
  /TARGET VARIABLES=activestb2_2A
  /FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
lifend

APPLY DICTIONARY
/ FROM BDH /SOURCE VARIABLES=cse_SP
/TARGET VARIABLES=cse_2A

```

```

/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

APPLY DICTIONARY
/FROM BDH /SOURCE VARIABLES=ocupac_SP
/TARGET VARIABLES=ocupac_2A
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

APPLY DICTIONARY
/FROM BDH /SOURCE VARIABLES=situprof_SP
/TARGET VARIABLES=situprof_2A
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

APPLY DICTIONARY
/FROM BDH /SOURCE VARIABLES=tipcont_SP
/TARGET VARIABLES=tipcont_2A
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

APPLY DICTIONARY
/FROM BDH /SOURCE VARIABLES=situac3_SP
/TARGET VARIABLES=situac3_2A
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

DATASET ACTIVATE ArchivoP .
SORT CASES BY idp (A) .
execute.

*-----,
ECHO ">>> comprobaciones 2º Adulto".

DATASET ACTIVATE ArchivoH .
FRECUENCIES VARIABLES=sexo_2A edad_2A edad_2A ned_2A naci_2A estcivil_2A
cse_2A ocupac_2A activestb1_2A activestb2_2A situprof_2A tipcont_2A situac3_2A /ORDER=ANALYSIS.

*---- activestb1_2A_m: Actividad del establecimiento laboral (1ª clasif) -----,
*----- (hasta el año 2009) Recodificación numérica -----,
DATASET ACTIVATE ArchivoH .
RECODE activestb1_2A (" ' '=SYSMIS) ('_NA'=-2) ('_ND'=0) ('A'+B'=1) ('C'+D+E'=2) ('F'=3) ('G'=4) ('H'=5) ('I'=6) ('J'=7)
('K'=8) ('L'=9) ('M'=10) ('N'=11) ('O'+P+Q'=12) INTO activestb1_2A_m.
if (SITUACT_SP ge 3 and SITUACT_SP le 7) activestb1_2A_m =-2 /* definido el la ECV solo para ocupados */.
if (gedad_2A=0) activestb1_2A_m =0 /* 0 ND (menor de 16 años) */.
execute.
APPLY DICTIONARY
/FROM ArchivoH /SOURCE VARIABLES=activestb1_SP_m
/TARGET VARIABLES=activestb1_2A_m
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

!if (!lepf lge 'EPF09') !then
*---- activestb2_2A_m: Actividad del establecimiento laboral (2ª clasif) -----,
*----- (a partir del año 2010) Recodificación numérica -----,
DATASET ACTIVATE ArchivoH .
RECODE activestb2_2A (" ' '=SYSMIS) ('_NA'=-2) ('_ND'=0) ('A'=1) ('B'=2) ('C'=3) ('D'=4) ('E'=5) ('F'=6) ('G'=7)
('H'=8) ('I'=9) ('J'=10) ('K'=11) ('L'=12) ('M'=13) ('N'=14) ('O'=15) ('P'=16)
('Q'=17) ('R'=18) ('S'=19) ('T'=20) ('U'=21) INTO activestb2_2A_m.
if (SITUACT_SP ge 3 and SITUACT_SP le 7) activestb2_2A_m =-2 /* definido el la ECV solo para ocupados */.
if (gedad_2A=0) activestb2_2A_m =0 /* 0 ND (menor de 16 años) */.
EXECUTE.
APPLY DICTIONARY
/FROM ArchivoH /SOURCE VARIABLES=activestb2_SP_m
/TARGET VARIABLES=activestb2_2A_m
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
!ifend

*=====,
** Carga de los Gastos del Hogar .
*=====,

!let !arch=!quote(!concat('FuentesDir\EPF\', !lepf, '\FichasG', !lepf, '.sav')).
GET FILE=!arch.
DATASET NAME Arch_Gastos WINDOW=FRONT.

DATASET ACTIVATE Arch_Gastos.
COMPUTE COD_GGASTO=TRUNC(CODIGO/1000).
EXECUTE.

DATASET DECLARE Gastos_Hogares.
SORT CASES BY NUMERO COD_GGASTO.
AGGREGATE
/OUTFILE='Gastos_Hogares'
/PRESORTED

```

330

* Eliminacion de atípicos-extremos.

*=====.

** Análisis de los Ingresos Monetarios Totales .

```
* EXAMINE VARIABLES=LINTOTALES
/PLOT BOXPLOT
/COMPARE GROUPS
/PERCENTILES(5,10,25,50,75,90,95, 99.8, 99.9) HAVERAGE
/STATISTICS DESCRIPTIVES EXTREME
/MISSING LISTWISE
/NOTOTAL.
```

```
* compute noextremo=not(Missig(LINTOTALES)) and (LINTOTALES>4.1572-3*(4.5101-4.1572)) and (LINTOTALES<4.5101+3*(4.5101-4.1572)).
```

```
* execute.
```

```
* use all.
* FILTER by noextremo.
* execute.
```

```
* EXAMINE VARIABLES=LINTOTALES
/PLOT BOXPLOT
/COMPARE GROUPS
/PERCENTILES(5,10,25,50,75,90,95, 99.8, 99.9) HAVERAGE
/STATISTICS DESCRIPTIVES EXTREME
/MISSING LISTWISE
/NOTOTAL.
```

** Análisis del índice de cobertura.

```
* EXAMINE VARIABLES=IC
/PLOT BOXPLOT
/COMPARE GROUPS
/PERCENTILES(5,10,25,50,75,90,95, 99.8, 99.9) HAVERAGE
/STATISTICS DESCRIPTIVES EXTREME
/MISSING LISTWISE
/NOTOTAL.
```

```
* compute LIC=lg10(IC).
```

```
* execute.
```

```
* EXAMINE VARIABLES=LIC
/PLOT BOXPLOT
/COMPARE GROUPS
/PERCENTILES(5,10,25,50,75,90,95,99.8, 99.9) HAVERAGE
/STATISTICS DESCRIPTIVES EXTREME
/MISSING LISTWISE
/NOTOTAL.
```

```
* compute noextremo=noextremo and not(Missig(LIC)) and (LIC>-0.0260-3*(0.2330-(-0.0260))) and (LIC<0.2330+3*(0.2330-(-0.0260))).
```

```
* execute.
```

```
* use all.
* FILTER by noextremo.
* execute.
```

```
* EXAMINE VARIABLES=LIC
/PLOT BOXPLOT
/COMPARE GROUPS
/PERCENTILES(5,10,25,50,75,90,95,99.9) HAVERAGE
/STATISTICS DESCRIPTIVES EXTREME
/MISSING LISTWISE
/NOTOTAL.
```

```
* EXAMINE VARIABLES=IC
/PLOT BOXPLOT
/COMPARE GROUPS
/PERCENTILES(5,10,25,50,75,90,95, 99.8, 99.9) HAVERAGE
/STATISTICS DESCRIPTIVES EXTREME
/MISSING LISTWISE
/NOTOTAL.
```

```
* EXAMINE VARIABLES=INTOTALES
/PLOT BOXPLOT
/COMPARE GROUPS
/PERCENTILES(5,10,25,50,75,90,95, 99.8, 99.9) HAVERAGE
/STATISTICS DESCRIPTIVES EXTREME
/MISSING LISTWISE
/NOTOTAL.
```

```

*** eliminacion de casos extremos .

* select if noextremo.
* execute.

*=====
**** correcciones por compatibilizar definiciones de "_NA" entre ECV, EPF y EFF.

DATASET ACTIVATE BDH .
if (tipcont_SP = 3) tipcont_SP = -2 .
if ((situac3_SP>=2 and situac3_SP<=6) or (situprof_SP=1 or situprof_SP=2 or situprof_SP=4)) tipcont_SP=-2 /* -2 No asalariado
(No Aplicable) */.
if (gedad_SP=0) tipcont_SP=0 /* 0 ND (menor de 16 años) */.
if (tipcont_2A = 3) tipcont_2A = -2 .
if ((situac3_2A>=2 and situac3_2A<=6) or (situprof_2A=1 or situprof_2A=2 or situprof_2A=4)) tipcont_2A=-2 /* -2 No
asalariado (No Aplicable) */.
if (gedad_2A=0) tipcont_2A=0 /* 0 ND (menor de 16 años) */.
execute.
ADD value labels tipcont_SP tipcont_2A -2 '_NA' 0 '_ND'.

if (situac3_SP>=2 and situac3_SP<=6) situprof_SP=-2 /* -2 No Aplicable - definido en la ECV sólo para ocupados*/.
if (gedad_SP=0) situprof_SP=0 /* 0 ND (menor de 16 años) */.
if (situac3_2A>=2 and situac3_2A<=6) situprof_2A=-2 /* -2 No Aplicable - definido en la ECV sólo para ocupados*/.
if (gedad_2A=0) situprof_2A=0 /* 0 ND (menor de 16 años) */.
execute.
ADD value labels situprof_SP situprof_2A -2 '_NA' 0 '_ND'.

if (situac3_SP>=2 and situac3_SP<=6) ocupac_SP=-2 /* -2 No Aplicable - definido en la ECV sólo para ocupados*/.
if (gedad_SP=0) ocupac_SP=0 /* 0 ND (menor de 16 años) */.
if (situac3_2A>=2 and situac3_2A<=6) ocupac_2A=-2 /* -2 No Aplicable - definido en la ECV sólo para ocupados*/.
if (gedad_2A=0) ocupac_2A=0 /* 0 ND (menor de 16 años) */.
execute.
ADD value labels ocupac_SP ocupac_2A -2 '_NA' 0 '_ND'.

if (situac3_SP>=2 and situac3_SP<=6) activestb1_SP_m=-2 /* -2 No Aplicable - definido en la ECV sólo para ocupados*/.
if (gedad_SP=0) activestb1_SP_m=0 /* 0 ND (menor de 16 años) */.
if (situac3_2A>=2 and situac3_2A<=6) activestb1_2A_m=-2 /* -2 No Aplicable - definido en la ECV sólo para
ocupados*/.
if (gedad_2A=0) activestb1_2A_m=0 /* 0 ND (menor de 16 años) */.
execute.
ADD value labels activestb1_SP_m activestb1_2A_m -2 '_NA' 0 '_ND'.

if (situac3_SP>=2 and situac3_SP<=6) activestb2_SP_m=-2 /* -2 No Aplicable - definido en la ECV sólo para ocupados*/.
if (gedad_SP=0) activestb2_SP_m=0 /* 0 ND (menor de 16 años) */.
if (situac3_2A>=2 and situac3_2A<=6) activestb2_2A_m=-2 /* -2 No Aplicable - definido en la ECV sólo para
ocupados*/.
if (gedad_2A=0) activestb2_2A_m=0 /* 0 ND (menor de 16 años) */.
execute.
ADD value labels activestb2_SP_m activestb2_2A_m -2 '_NA' 0 '_ND'.

*=====
* Grabación del Archivo_Match.
*=====

DATASET ACTIVATE ArchivoH .
!!let !arch=!quote(!concat('ArchivoH_', !epf, '_para_matching.sav')).
save OUTFILE=!arch /COMPRESSED .
EXECUTE .

*=====
* Salida .
*=====

!!let !arch=!quote(!concat('ArchivoH_', !epf, '_para_matching.spv')).
OUTPUT SAVE NAME=salida OUTFILE=!arch LOCK=NO.
OUTPUT CLOSE salida.

New file.
dataset close all.

!enddefine.

```


Anexo 1.c. Programa que prepara los Datos de Hogares de las EFF para matching

```

=====
*==>>> 03a_Prepara_Datos_EFF_H_para_matching.sps
=====

define !Prepara_Datos_EFF_H_para_matching (eff = !TOKENS(1) ).
!!let !ano=!concat('20', !substr(!eff,4)).

New File.
dataset close ALL.

OUTPUT NEW NAME =salida.
OUTPUT ACTIVATE salida.

cd 'WorkDir'.

=====
* Carga de la BDH_Match en memoria .
=====
GET FILE='F_BDHDir\BDH_para_matching.sav'.
DATASET NAME BDH.

=====
* Carga de los datos del Hogar, ArchivoH, de la EFF .
=====
!!let !arch=!quote(!concat( 'ArchivoH_', !eff, '.sav')).
GET FILE=!arch.
DATASET NAME ArchivoH.
dataset activate ArchivoH.

=====
* Carga de los datos básicos del Hogar, H, de la EFF .
=====

*----- idh: Identificador del Hogar -----,
compute idh=idhogar.
execute.
Apply Dictionary
  /FROM BDH /SOURCE VARIABLES=idh
  /TARGET VARIABLES=idh
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
Formats idh (F10.0).

*----- w: Factor de elevación del hogar -----,
compute w=fehogar.
execute.
Apply Dictionary
  /FROM BDH /SOURCE VARIABLES=w
  /TARGET VARIABLES=w
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

=====
* Composicion y Tipología del Hogar .
=====

*----- nmh: Número de Miembros del Hogar -----,
  !if (!eff='EFF08') !then
compute nmh=p1+p1_11.
  !ifend
  !if (!eff='EFF11') !then
compute nmh=p1.
  !ifend
execute.
Apply Dictionary
  /FROM BDH /SOURCE VARIABLES=nmh
  /TARGET VARIABLES=nmh
  /FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*----- nmh16: Número de miembros del Hogar con 16 años o más -----,
compute nmh16=0.
  !if (!eff='EFF08') !then
do repeat v=p1_2d_1 p1_2d_2 p1_2d_3 p1_2d_4 p1_2d_5 p1_2d_6 p1_2d_7 p1_2d_8 p1_2d_9
  p1_12_1 p1_12_2 p1_12_3 p1_12_4 p1_12_5 .
  do if ( not(missing(v)) and (v>=16) ).
    compute nmh16=nmh16+1.
  
```

```

end if.
end repeat.
!ifend
!if (leff='EFF11') !then
do repeat v=p1_2d_1 p1_2d_2 p1_2d_3 p1_2d_4 p1_2d_5 p1_2d_6 p1_2d_7 p1_2d_8 p1_2d_9 .
do if ( not(missing(v)) and (v>=16) ).
compute nmh16=nmh16+1.
end if.
end repeat.
!ifend
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=nmh16
/TARGET VARIABLES=nmh16
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*----- nmh14: Número de miembros del Hogar con 14 años o más -----
compute nmh14=0.
!if (leff='EFF08') !then
do repeat v=p1_2d_1 p1_2d_2 p1_2d_3 p1_2d_4 p1_2d_5 p1_2d_6 p1_2d_7 p1_2d_8 p1_2d_9
p1_12_1 p1_12_2 p1_12_3 p1_12_4 p1_12_5 .
do if ( not(missing(v)) and (v>=14) ).
compute nmh14=nmh14+1.
end if.
end repeat.
!ifend
!if (leff='EFF11') !then
do repeat v=p1_2d_1 p1_2d_2 p1_2d_3 p1_2d_4 p1_2d_5 p1_2d_6 p1_2d_7 p1_2d_8 p1_2d_9 .
do if ( not(missing(v)) and (v>=14) ).
compute nmh14=nmh14+1.
end if.
end repeat.
!ifend
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=nmh14
/TARGET VARIABLES=nmh14
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

=====
* Variables específicas del Hogar para el maching.
=====

*----- MATCH - REGION_m: Comunidad Autónoma (No disponible) -----
*----- MATCH - HAB_m : Habitat de la vivienda (No disponible) -----
*----- MATCH - TIPOVIV_m: Tipo de vivienda (No disponible) -----

*----- MATCH - REGTENEVIV3_m: Régimen de tenencia de la vivienda -----
compute Regteneviv3_m=$SYSMIS.
if (p2_1=2 and (p2_7=2 or p2_8=2)) Regteneviv3_m=1.
if (p2_1=2 and (p2_7=1 and p2_8=1)) Regteneviv3_m=2.
if (p2_1=1) Regteneviv3_m=3.
if (p2_1=3) Regteneviv3_m=4.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=Regteneviv3_m
/TARGET VARIABLES=Regteneviv3_m
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*----- MATCH - NUMHABIT_m: Número de habitaciones de la vivienda -----

=====
* Principales datos resúmenes de Composicion del Hogar .
=====

*>>> comprobación de variables de Situación laboral (año en curso).
do repeat v1 = p6_1c1_1 p6_1c1_2 p6_1c1_3 p6_1c1_4 p6_1c1_5 p6_1c1_6 p6_1c1_7 p6_1c1_8 p6_1c1_9
/v2 = p6_1c2_1 p6_1c2_2 p6_1c2_3 p6_1c2_4 p6_1c2_5 p6_1c2_6 p6_1c2_7 p6_1c2_8 p6_1c2_9
/v3 = p6_1c3_1 p6_1c3_2 p6_1c3_3 p6_1c3_4 p6_1c3_5 p6_1c3_6 p6_1c3_7 p6_1c3_8 p6_1c3_9
/v4 = p6_1c4_1 p6_1c4_2 p6_1c4_3 p6_1c4_4 p6_1c4_5 p6_1c4_6 p6_1c4_7 p6_1c4_8 p6_1c4_9
/v5 = p6_1c5_1 p6_1c5_2 p6_1c5_3 p6_1c5_4 p6_1c5_5 p6_1c5_6 p6_1c5_7 p6_1c5_8 p6_1c5_9
/v6 = p6_1c6_1 p6_1c6_2 p6_1c6_3 p6_1c6_4 p6_1c6_5 p6_1c6_6 p6_1c6_7 p6_1c6_8 p6_1c6_9
/v7 = p6_1c7_1 p6_1c7_2 p6_1c7_3 p6_1c7_4 p6_1c7_5 p6_1c7_6 p6_1c7_7 p6_1c7_8 p6_1c7_9
/v8 = p6_1c8_1 p6_1c8_2 p6_1c8_3 p6_1c8_4 p6_1c8_5 p6_1c8_6 p6_1c8_7 p6_1c8_8 p6_1c8_9
/sitac= sitac1 sitac2 sitac3 sitac4 sitac5 sitac6 sitac7 sitac8 sitac9
/diff= d1 d2 d3 d4 d5 d6 d7 d8 d9 .
do if (v1=1 or v2=1).
compute sitac=1 /* 1- ocupado */.

```

```

else if (v4=1 or v5=1).
  compute sitac=4 /* 4- retirado */ .
else if (v6=1).
  compute sitac=3 /* 3- estudiante */ .
else if (v3=1).
  compute sitac=2 /* 2- desempleado */ .
else if (v7=1).
  compute sitac=5 /* 5 - labores del Hogar */ .
else if (v8=1).
  compute sitac=6 /* 6 - otra inactividad */ .
end if.
compute diff=1-sum(v1,v2,v3,v4,v5,v6,v7,v8).
end repeat.
execute.
FRECUENCIAS VARIABLES=d1 d2 d3 d4 d5 d6 d7 d8 d9.
delete variables d1 d2 d3 d4 d5 d6 d7 d8 d9.

*>>> comprobación de variables de Situación laboral (año anterior).
do repeat v1 = p6_63c1_1 p6_63c1_2 p6_63c1_3 p6_63c1_4 p6_63c1_5 p6_63c1_6 p6_63c1_7 p6_63c1_8 p6_63c1_9
  / v2 = p6_63c2_1 p6_63c2_2 p6_63c2_3 p6_63c2_4 p6_63c2_5 p6_63c2_6 p6_63c2_7 p6_63c2_8 p6_63c2_9
  / v3 = p6_63c3_1 p6_63c3_2 p6_63c3_3 p6_63c3_4 p6_63c3_5 p6_63c3_6 p6_63c3_7 p6_63c3_8 p6_63c3_9
  / v4 = p6_63c4_1 p6_63c4_2 p6_63c4_3 p6_63c4_4 p6_63c4_5 p6_63c4_6 p6_63c4_7 p6_63c4_8 p6_63c4_9
  / v5 = p6_63c5_1 p6_63c5_2 p6_63c5_3 p6_63c5_4 p6_63c5_5 p6_63c5_6 p6_63c5_7 p6_63c5_8 p6_63c5_9
  / v6 = p6_63c6_1 p6_63c6_2 p6_63c6_3 p6_63c6_4 p6_63c6_5 p6_63c6_6 p6_63c6_7 p6_63c6_8 p6_63c6_9
  / v7 = p6_63c7_1 p6_63c7_2 p6_63c7_3 p6_63c7_4 p6_63c7_5 p6_63c7_6 p6_63c7_7 p6_63c7_8 p6_63c7_9
  / v8 = p6_63c8_1 p6_63c8_2 p6_63c8_3 p6_63c8_4 p6_63c8_5 p6_63c8_6 p6_63c8_7 p6_63c8_8 p6_63c8_9
  / sitac= sitac1AA sitac2AA sitac3AA sitac4AA sitac5AA sitac6AA sitac7AA sitac8AA sitac9AA
  / diff= d1 d2 d3 d4 d5 d6 d7 d8 d9 .
do if (v1=1 or v2=1).
  compute sitac=1 /* 1- ocupado */ .
else if (v4=1 or v5=1).
  compute sitac=4 /* 4- retirado */ .
else if (v6=1).
  compute sitac=3 /* 3- estudiante */ .
else if (v3=1).
  compute sitac=2 /* 2- desempleado */ .
else if (v7=1).
  compute sitac=5 /* 5 - labores del Hogar */ .
else if (v8=1).
  compute sitac=6 /* 6 - otra inactividad */ .
end if.
compute diff=1-sum(v1,v2,v3,v4,v5,v6,v7,v8).
end repeat.
execute.
FRECUENCIAS VARIABLES=d1 d2 d3 d4 d5 d6 d7 d8 d9.
delete variables d1 d2 d3 d4 d5 d6 d7 d8 d9.

* ----- n_dep: número de dependientes del hogar (según la ECV) -----
*----- n_adult: número personas de 18 o más años y que no son dependientes -.
compute n_dep=0.
compute n_adult=0.
do repeat v=p1_2d_1 p1_2d_2 p1_2d_3 p1_2d_4 p1_2d_5 p1_2d_6 p1_2d_7 p1_2d_8 p1_2d_9
  / sitac= sitac1 sitac2 sitac3 sitac4 sitac5 sitac6 sitac7 sitac8 sitac9 .
do if (not(missing(v))).
do if (v<18).
  compute n_dep=n_dep+1.
else if ( (v<25) and (sitac=3 or sitac=4 or sitac=5 or sitac=6) ) .
  compute n_dep=n_dep+1.
else .
  compute n_adult=n_adult+1.
end if.
end if.
end repeat.
execute.
Apply Dictionary
/ FROM BDH /SOURCE VARIABLES=n_dep
/TARGET VARIABLES=n_dep
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

* ----- n_hombre: número de hombres del hogar -----
compute n_hombre=0.
do repeat v=p1_1_1 p1_1_2 p1_1_3 p1_1_4 p1_1_5 p1_1_6 p1_1_7 p1_1_8 p1_1_9 .
do if (v=1).
  compute n_hombre=n_hombre+1.
end if.
end repeat.
execute.
Apply Dictionary
/ FROM BDH /SOURCE VARIABLES=n_hombre
/TARGET VARIABLES=n_hombre
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

```

```

* ----- n_mujer: número de mujeres del hogar -----,
compute n_mujer=0.
do repeat v=p1_1_1 p1_1_2 p1_1_3 p1_1_4 p1_1_5 p1_1_6 p1_1_7 p1_1_8 p1_1_9 .
do if (v=2).
  compute n_mujer=n_mujer+1.
end if.
end repeat.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=n_mujer
/TARGET VARIABLES=n_mujer
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

```

```

* ----- n_soltero: número de solteros del hogar -----,
compute n_soltero=0.
do repeat v=p1_4_1 p1_4_2 p1_4_3 p1_4_4 p1_4_5 p1_4_6 p1_4_7 p1_4_8 p1_4_9 .
do if (v=1).
  compute n_soltero=n_soltero+1.
end if.
end repeat.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=n_soltero
/TARGET VARIABLES=n_soltero
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

```

```

* ----- n_casado: número de casados del hogar (o parejas de hecho) -----,
compute n_casado=0.
do repeat v=p1_4_1 p1_4_2 p1_4_3 p1_4_4 p1_4_5 p1_4_6 p1_4_7 p1_4_8 p1_4_9 .
do if (v=2 or v=3).
  compute n_casado=n_casado+1.
end if.
end repeat.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=n_casado
/TARGET VARIABLES=n_casado
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

```

```

* ----- n_separado: número de separados del hogar -----,
compute n_separado=0.
do repeat v=p1_4_1 p1_4_2 p1_4_3 p1_4_4 p1_4_5 p1_4_6 p1_4_7 p1_4_8 p1_4_9 .
do if (v=4).
  compute n_separado=n_separado+1.
end if.
end repeat.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=n_separado
/TARGET VARIABLES=n_separado
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

```

```

* ----- n_viudo: número de viudos del hogar -----,
compute n_viudo=0.
do repeat v=p1_4_1 p1_4_2 p1_4_3 p1_4_4 p1_4_5 p1_4_6 p1_4_7 p1_4_8 p1_4_9 .
do if (v=6).
  compute n_viudo=n_viudo+1.
end if.
end repeat.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=n_viudo
/TARGET VARIABLES=n_viudo
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

```

```

* ----- n_divorciado: número de divorciados del hogar -----,
compute n_divorciado=0.
do repeat v=p1_4_1 p1_4_2 p1_4_3 p1_4_4 p1_4_5 p1_4_6 p1_4_7 p1_4_8 p1_4_9 .
do if (v=5).
  compute n_divorciado=n_divorciado+1.
end if.
end repeat.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=n_divorciado
/TARGET VARIABLES=n_divorciado
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

```

```

* ----- n_espanola: número de españoles del hogar (no disponible) -----,

```

```

* ----- n_exUE24: número de miembros extranjeros del resto de la UE-25 del hogar (no disponible)-.

* ----- n_exResE: número de miembros extranjeros del resto de Europa del hogar (no disponible) --.

* ----- n_exResM: número de extranjeros del resto del mundo en el hogar (no disponible) -----.

* ----- n_sinestud: número de miembros sin estudios del hogar -----.
compute n_sinestud=0.
do repeat v=p1_5_1 p1_5_2 p1_5_3 p1_5_4 p1_5_5 p1_5_6 p1_5_7 p1_5_8 p1_5_9 .
do if (v=1).
  compute n_sinestud=n_sinestud+1.
end if.
end repeat.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=n_sinestud
/TARGET VARIABLES=n_sinestud
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

* ----- n_eduprimaria: número de miembros sin estudios o con Educación Primaria del hogar -----.
compute n_eduprimaria=0.
do repeat v=p1_5_1 p1_5_2 p1_5_3 p1_5_4 p1_5_5 p1_5_6 p1_5_7 p1_5_8 p1_5_9 .
do if (v=2).
  compute n_eduprimaria=n_eduprimaria+1.
end if.
end repeat.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=n_eduprimaria
/TARGET VARIABLES=n_eduprimaria
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

* ----- n_edusec1: número de miembros con Educación Secundaria de 1ª etapa del hogar -----.
compute n_edusec1=0.
do repeat v=p1_5_1 p1_5_2 p1_5_3 p1_5_4 p1_5_5 p1_5_6 p1_5_7 p1_5_8 p1_5_9 .
do if (v=3 or v=4).
  compute n_edusec1=n_edusec1+1.
end if.
end repeat.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=n_edusec1
/TARGET VARIABLES=n_edusec1
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

* ----- n_edusec2: número de miembros con Educación Secundaria de 2ª etapa del hogar -----.
compute n_edusec2=0.
do repeat v=p1_5_1 p1_5_2 p1_5_3 p1_5_4 p1_5_5 p1_5_6 p1_5_7 p1_5_8 p1_5_9 .
do if (v=5 or v=6).
  compute n_edusec2=n_edusec2+1.
end if.
end repeat.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=n_edusec2
/TARGET VARIABLES=n_edusec2
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

* ----- n_edusuperior: número de miembros con Educación Superior del hogar -----.
compute n_edusuperior=0.
do repeat v=p1_5_1 p1_5_2 p1_5_3 p1_5_4 p1_5_5 p1_5_6 p1_5_7 p1_5_8 p1_5_9 .
do if (v=7 or v=8 or v=9 or v=10 or v=11 or v=12).
  compute n_edusuperior=n_edusuperior+1.
end if.
end repeat.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=n_edusuperior
/TARGET VARIABLES=n_edusuperior
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

* ----- n_ocupado: número de ocupados del hogar -----.
compute n_ocupado=0.
do repeat sitac= sitac1 sitac2 sitac3 sitac4 sitac5 sitac6 sitac7 sitac8 sitac9 .
do if (sitac=1).
  compute n_ocupado=n_ocupado+1.
end if.
end repeat.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=n_ocupado

```

```

/TARGET VARIABLES=n_ocupado
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

* ----- n_desempleado: número de activos desempleados del hogar -----
compute n_desempleado=0.
do repeat sitac= sitac1 sitac2 sitac3 sitac4 sitac5 sitac6 sitac7 sitac8 sitac9 .
do if (sitac=2).
compute n_desempleado=n_desempleado+1.
end if.
end repeat.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=n_desempleado
/TARGET VARIABLES=n_desempleado
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

* ----- n_estudiante: número de inactivos estudiantes del hogar -----
compute n_estudiante=0.
do repeat sitac= sitac1 sitac2 sitac3 sitac4 sitac5 sitac6 sitac7 sitac8 sitac9 .
do if (sitac=3).
compute n_estudiante=n_estudiante+1.
end if.
end repeat.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=n_estudiante
/TARGET VARIABLES=n_estudiante
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

* ----- n_retirado: número de inactivos jubilados o retirados del hogar -----
compute n_retirado=0.
do repeat sitac= sitac1 sitac2 sitac3 sitac4 sitac5 sitac6 sitac7 sitac8 sitac9 .
do if (sitac=4).
compute n_retirado=n_retirado+1.
end if.
end repeat.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=n_retirado
/TARGET VARIABLES=n_retirado
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

* ----- n_labhogar: número inactivos dedicados a labores del hogar -----
compute n_labhogar=0.
do repeat sitac= sitac1 sitac2 sitac3 sitac4 sitac5 sitac6 sitac7 sitac8 sitac9 .
do if (sitac=5).
compute n_labhogar=n_labhogar+1.
end if.
end repeat.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=n_labhogar
/TARGET VARIABLES=n_labhogar
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

* ----- n_otrinactivo: número de inactivos del hogar -----
compute n_otrinactivo=0.
do repeat sitac= sitac1 sitac2 sitac3 sitac4 sitac5 sitac6 sitac7 sitac8 sitac9 .
do if (sitac=6).
compute n_otrinactivo=n_otrinactivo+1.
end if.
end repeat.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=n_otrinactivo
/TARGET VARIABLES=n_otrinactivo
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

* ---- comprobaciones de variables resúmenes de hogares -----

DATASET ACTIVATE ArchivoH.
compute diff_dep = Min(9,nmh)-(n_dep+n_adult).
compute diff_sexo = Min(9,nmh)-(n_hombre+n_mujer).
compute diff_estcivil = Min(9,nmh16)-(n_soltero+n_casado+n_viudo+n_separado+n_divorciado).
compute diff_nestud = Min(9,nmh16)-(n_sinestud+n_eduprimaria+n_edusec1+n_edusec2+n_edusuperior).
compute diff_situac = Min(9,nmh16)-(n_ocupado+n_desempleado+n_estudiante+n_retirado+n_labhogar+n_otrinactivo) .
execute.
FRECUENCIAS VARIABLES=diff_dep diff_sexo diff_estcivil diff_nestud diff_situac .
FRECUENCIAS VARIABLES=n_hombre n_mujer n_soltero n_casado n_viudo n_separado n_divorciado
n_sinestud n_eduprimaria n_edusec1 n_edusec2 n_edusuperior n_ocupado n_desempleado n_estudiante n_retirado
n_labhogar n_otrinactivo

```

```

/ORDER=ANALYSIS.
DELETE VARIABLES diff_dep diff_sexo diff_estcivil diff_nestud diff_situac.

*----- MATCH - TIPHOGAR_m: Tipo de hogar -----,

compute n_desc=nmh-(n_dep+n_adult) /* missings para n_dep */.
compute max_edad=-999.
do repeat v=p1_2d_1 p1_2d_2 p1_2d_3 p1_2d_4 p1_2d_5 p1_2d_6 p1_2d_7 p1_2d_8 p1_2d_9 .
  if (not(missing(v))) max_edad=max(max_edad,v).
end repeat.
execute.

compute TIPHOGAR_m=$SYSMIS.
Do if (nmh=1 and n_desc=0 and n_adult=1).
  if (max_edad < 30) TIPHOGAR_m=1.
  if (30<=max_edad and max_edad<65) TIPHOGAR_m=2.
  if (max_edad >=65) TIPHOGAR_m=3.
Else if (nmh=1 and n_desc=0 and n_dep=1).
  compute TIPHOGAR_m=14.
Else if (nmh=2) and (n_desc=0) .
  if (n_adult=2 and n_dep=0 and max_edad>=65) TIPHOGAR_m=7.
  if (n_adult=2 and n_dep=0 and max_edad <65) TIPHOGAR_m=8.
  if (n_adult=1) and (n_dep=1) TIPHOGAR_m=10.
  if (n_adult=0) and (n_dep=2) TIPHOGAR_m=14.
Else if (nmh>2) and (n_desc=0).
  do if (n_dep=0).
    compute TIPHOGAR_m=9.
  Else if (n_dep>0) .
    if (n_adult=0) TIPHOGAR_m=14.
    if (n_adult=1) TIPHOGAR_m=10.
    if (n_adult=2) and (n_dep=1) TIPHOGAR_m=11.
    if (n_adult=2) and (n_dep=2) TIPHOGAR_m=12.
    if (n_adult=2) and (n_dep>=3) TIPHOGAR_m=13.
    if (n_adult>2) TIPHOGAR_m=14.
  End if.
End if.
execute.
Delete Variables max_edad n_desc.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=TIPHOGAR_m
/TARGET VARIABLES=TIPHOGAR_m
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*=====,
** Variables sobre los Ingresos del Hogar .
*=====,

*----- ano: Año de referencia de los ingresos -----,
Numeric ano (F4.0) .
Variable Labels ano 'Año de Referencia de los ingresos'.
  !if (!eff='EFF08') !then
Value labels / ano
  2008 'año 2008'.
compute ano=2008.
  !ifend
  !if (!eff='EFF11') !then
Value labels / ano
  2011 'año 2011'.
compute ano=2011.
  !ifend
execute.

*----- um: Unidad de medida de los ingresos -----,
String um (A8) .
Variable Labels um 'Unidad de medida de los ingresos'.
  !if (!eff='EFF08') !then
Value labels / um
  '€' 'euros corrientes' .
compute um='€'.
  !ifend
  !if (!eff='EFF11') !then
Value labels / um
  '€' 'euros corrientes' .
compute um='€'.
  !ifend
execute.

*----- RMCP_B : Renta Monetaria por Cuenta Propia (bruta) -----,
Numeric RMCP_B (F10.2).

```

```

compute RMCP_B=0.
Missing Values p6_102_1_1 p6_102_2_1 p6_102_3_1 p6_102_4_1 p6_102_5_1 p6_102_6_1 p6_102_7_1 p6_102_8_1 p6_102_9_1
p6_102_1_2 p6_102_2_2 p6_102_3_2 p6_102_4_2 p6_102_5_2 p6_102_6_2 p6_102_7_2 p6_102_8_2 p6_102_9_2
p6_102_1_3 p6_102_2_3 p6_102_3_3 p6_102_4_3 p6_102_5_3 p6_102_6_3 p6_102_7_3 p6_102_8_3 p6_102_9_3
p6_104_1_1 p6_104_2_1 p6_104_3_1 p6_104_4_1 p6_104_5_1 p6_104_6_1 p6_104_7_1 p6_104_8_1 p6_104_9_1
p6_104_1_2 p6_104_2_2 p6_104_3_2 p6_104_4_2 p6_104_5_2 p6_104_6_2 p6_104_7_2 p6_104_8_2 p6_104_9_2
p6_104_1_3 p6_104_2_3 p6_104_3_3 p6_104_4_3 p6_104_5_3 p6_104_6_3 p6_104_7_3 p6_104_8_3 p6_104_9_3 (-1, -
2).
do repeat AAAA = p6_102_1_1 p6_102_2_1 p6_102_3_1 p6_102_4_1 p6_102_5_1 p6_102_6_1 p6_102_7_1 p6_102_8_1
p6_102_9_1
p6_102_1_2 p6_102_2_2 p6_102_3_2 p6_102_4_2 p6_102_5_2 p6_102_6_2 p6_102_7_2 p6_102_8_2 p6_102_9_2
p6_102_1_3 p6_102_2_3 p6_102_3_3 p6_102_4_3 p6_102_5_3 p6_102_6_3 p6_102_7_3 p6_102_8_3 p6_102_9_3
/ BBBB = p6_104_1_1 p6_104_2_1 p6_104_3_1 p6_104_4_1 p6_104_5_1 p6_104_6_1 p6_104_7_1 p6_104_8_1 p6_104_9_1
p6_104_1_2 p6_104_2_2 p6_104_3_2 p6_104_4_2 p6_104_5_2 p6_104_6_2 p6_104_7_2 p6_104_8_2 p6_104_9_2
p6_104_1_3 p6_104_2_3 p6_104_3_3 p6_104_4_3 p6_104_5_3 p6_104_6_3 p6_104_7_3 p6_104_8_3 p6_104_9_3.
if (not(missing(AAAA))) RMCP_B=RMCP_B+12*AAAA.
if (not(missing(BBBB))) RMCP_B=RMCP_B+BBBB.
end repeat.
Missing Values p4_112_1 p4_112_2 p4_112_3 p4_112_4 p4_112_5 p4_112_6
p4_112b_1 p4_112b_2 p4_112b_3 p4_112b_4 p4_112b_5 p4_112b_6 (-1, -2).
do repeat AAA = p4_112_1 p4_112_2 p4_112_3 p4_112_4 p4_112_5 p4_112_6
/ BBB = p4_112b_1 p4_112b_2 p4_112b_3 p4_112b_4 p4_112b_5 p4_112b_6 .
if (not(missing(AAA))) RMCP_B=RMCP_B+AAA.
if (not(missing(BBB))) RMCP_B=RMCP_B-BBB.
end repeat.
execute.
Variable Labels RMCP_B 'Renta Mixta Monetaria Bruta del Hogar'.
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=RMCP_B ngrupos=10 pi=sí .

*----- RNMCP_B: Renta No Monetaria por Cuenta Propia (se recogen neta) -----
Numeric RNMCP_B (F10.2).
compute RNMCP_B=0.
Missing Values P6_60f (-1, -2).
if (not(missing(P6_60f)) and (P6_60es1=2 or P6_60es2=0)) RNMCP_B=RNMCP_B+4*P6_60f.
execute .
Variable Labels RNMCP_B 'Renta Mixta no Monetaria Bruta del Hogar' .
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=RNMCP_B ngrupos=10 pi=sí .

*----- RMIX_B: RENTAS MIXTAS (brutas) -----
Numeric RMIX_B (F10.2).
compute RMIX_B=RMCP_B+RNMCP_B.
execute .
Variable Labels RMIX_B 'RENTAS MIXTAS (Brutas) del Hogar' .
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=RMIX_B ngrupos=10 pi=sí .

Numeric RMIX_BAA (F10.2).
compute RMIX_BAA=0.
Missing Values p6_72_1 p6_72_2 p6_72_3 p6_72_4 p6_72_5 p6_72_6 p6_72_7 p6_72_8 p6_72_9 (-1, -2).
do repeat AAA = p6_72_1 p6_72_2 p6_72_3 p6_72_4 p6_72_5 p6_72_6 p6_72_7 p6_72_8 p6_72_9 .
if (not(missing(AAA))) RMIX_BAA=RMIX_BAA+AAA.
end repeat.
execute.
Variable Labels RMIX_BAA 'Rentas Trabajo por Cuenta Propia (2010)'.

*----- RMCA_B: Renta monetaria o cuasimonetaria de asalariados (bruta) -----
Numeric RMCA_B (F10.2).
compute RMCA_B=0.
Missing Values p6_14_1_1 p6_14_2_1 p6_14_3_1 p6_14_4_1 p6_14_5_1 p6_14_6_1 p6_14_7_1 p6_14_8_1 p6_14_9_1
p6_14_1_2 p6_14_2_2 p6_14_3_2 p6_14_4_2 p6_14_5_2 p6_14_6_2 p6_14_7_2 p6_14_8_2 p6_14_9_2
p6_14_1_3 p6_14_2_3 p6_14_3_3 p6_14_4_3 p6_14_5_3 p6_14_6_3 p6_14_7_3 p6_14_8_3 p6_14_9_3
p6_14b_1_1 p6_14b_2_1 p6_14b_3_1 p6_14b_4_1 p6_14b_5_1 p6_14b_6_1 p6_14b_7_1 p6_14b_8_1 p6_14b_9_1
p6_14b_1_2 p6_14b_2_2 p6_14b_3_2 p6_14b_4_2 p6_14b_5_2 p6_14b_6_2 p6_14b_7_2 p6_14b_8_2 p6_14b_9_2
p6_14b_1_3 p6_14b_2_3 p6_14b_3_3 p6_14b_4_3 p6_14b_5_3 p6_14b_6_3 p6_14b_7_3 p6_14b_8_3 p6_14b_9_3 (-1,
-2).
do repeat AAA = p6_14_1_1 p6_14_2_1 p6_14_3_1 p6_14_4_1 p6_14_5_1 p6_14_6_1 p6_14_7_1 p6_14_8_1 p6_14_9_1
p6_14_1_2 p6_14_2_2 p6_14_3_2 p6_14_4_2 p6_14_5_2 p6_14_6_2 p6_14_7_2 p6_14_8_2 p6_14_9_2
p6_14_1_3 p6_14_2_3 p6_14_3_3 p6_14_4_3 p6_14_5_3 p6_14_6_3 p6_14_7_3 p6_14_8_3 p6_14_9_3
/ BBBB = p6_14b_1_1 p6_14b_2_1 p6_14b_3_1 p6_14b_4_1 p6_14b_5_1 p6_14b_6_1 p6_14b_7_1 p6_14b_8_1 p6_14b_9_1
p6_14b_1_2 p6_14b_2_2 p6_14b_3_2 p6_14b_4_2 p6_14b_5_2 p6_14b_6_2 p6_14b_7_2 p6_14b_8_2 p6_14b_9_2
p6_14b_1_3 p6_14b_2_3 p6_14b_3_3 p6_14b_4_3 p6_14b_5_3 p6_14b_6_3 p6_14b_7_3 p6_14b_8_3 p6_14b_9_3.
if (not(missing(AAA))) RMCA_B=RMCA_B+12*AAA.
if (not(missing(BBBB))) RMCA_B=RMCA_B-12*BBBB.
end repeat.
Missing Values P6_60c P6_60b (-1, -2).
if (not(missing(P6_60c))) RMCA_B=RMCA_B+12*P6_60c.
if (not(missing(P6_60b))) RMCA_B=RMCA_B-12*P6_60b.
execute .
Variable Labels RMCA_B 'Renta bruta monetaria o cuasimonetaria de asalariados del Hogar'.
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=RMCA_B ngrupos=10 pi=sí .

Numeric RMCA_BAA (F10.2).

```



```

compute RMCA_BAA=0.
Missing Values p6_64_1 p6_64_2 p6_64_3 p6_64_4 p6_64_5 p6_64_6 p6_64_7 p6_64_8 p6_64_9 (-1, -2).
do repeat AAA = p6_64_1 p6_64_2 p6_64_3 p6_64_4 p6_64_5 p6_64_6 p6_64_7 p6_64_8 p6_64_9.
    if (not(missing(AAA))) RMCA_BAA=RMCA_BAA+AAA.
end repeat.
missing Values P6_76b (-1, -2).
if (not(missing(P6_76b))) RMCA_BAA=RMCA_BAA+P6_76b .
execute.
Variable Labels RMCA_BAA 'Renta bruta monetaria o cuasimonetaria de asalariados del Hogar (2010)'.

*----- RNMCA_B: Renta bruta no monetaria de asalariados -----
Numeric RNMCA_B (F10.2).
compute RNMCA_B=0.
Missing Values p6_16_1_1 p6_16_2_1 p6_16_3_1 p6_16_4_1 p6_16_5_1 p6_16_6_1 p6_16_7_1 p6_16_8_1 p6_16_9_1
    p6_16_1_2 p6_16_2_2 p6_16_3_2 p6_16_4_2 p6_16_5_2 p6_16_6_2 p6_16_7_2 p6_16_8_2 p6_16_9_2
    p6_16_1_3 p6_16_2_3 p6_16_3_3 p6_16_4_3 p6_16_5_3 p6_16_6_3 p6_16_7_3 p6_16_8_3 p6_16_9_3 (-1, -2).
do repeat AAA = p6_16_1_1 p6_16_2_1 p6_16_3_1 p6_16_4_1 p6_16_5_1 p6_16_6_1 p6_16_7_1 p6_16_8_1 p6_16_9_1
    p6_16_1_2 p6_16_2_2 p6_16_3_2 p6_16_4_2 p6_16_5_2 p6_16_6_2 p6_16_7_2 p6_16_8_2 p6_16_9_2
    p6_16_1_3 p6_16_2_3 p6_16_3_3 p6_16_4_3 p6_16_5_3 p6_16_6_3 p6_16_7_3 p6_16_8_3 p6_16_9_3.
    if (not(missing(AAA))) RNMCA_B=RNMCA_B+12*AAA.
end repeat.
execute .
Variable Labels RNMCA_B 'Renta bruta no monetaria de asalariados del Hogar'.
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=RNMCA_B ngrupos=10 pi=sí .

Numeric RNMCA_BAA (F10.2).
compute RNMCA_BAA=0.
Missing Values p6_66_1 p6_66_2 p6_66_3 p6_66_4 p6_66_5 p6_66_6 p6_66_7 p6_66_8 p6_66_9 (-1, -2).
do repeat AAA = p6_66_1 p6_66_2 p6_66_3 p6_66_4 p6_66_5 p6_66_6 p6_66_7 p6_66_8 p6_66_9.
    if (not(missing(AAA))) RNMCA_BAA=RNMCA_BAA+AAA.
end repeat.
execute.
Variable Labels RNMCA_BAA 'Rentas salariales No Monetaria del Hogar (2010)'.

*----- RNM2CA_B: Renta bruta no monetaria de asalariados (coche de empresa) -----
***** >>>> (No se recogen en le encuesta) <<<< *****
Numeric RNM2CA_B (F10.2).
compute RNM2CA_B=$SYSMIS.
execute .
Variable Labels RNM2CA_B 'Renta bruta no monetaria de asalariados del Hogar (coche de empresa)'.

*----- CSEFEMP: Cotizaciones sociales efectivas a cargo del empleador -----
***** >>>> (No se recogen en le encuesta) <<<< *****
Numeric CSEFEMP (F10.2).
compute CSEFEMP=$SYSMIS.
execute .
Variable Labels CSEFEMP 'Cotizaciones sociales efectivas a cargo del empleador'.

*----- RASAL_B: RENTAS SALARIALES (Brutas) -----
Numeric RASAL_B (F10.2).
Do if (not(missing(CSEFEMP))).
    compute RASAL_B=RMCA_B+CSEFEMP+RNMCA_B.
else.
    compute RASAL_B=RMCA_B+RNMCA_B.
end if.
execute .
Variable Labels RASAL_B 'RENTAS SALARIALES (Brutas) del Hogar'.

Numeric RASAL_BAA (F10.2).
compute RASAL_BAA=RMCA_BAA+RNMCA_BAA.
execute .
Variable Labels RASAL_BAA 'RENTAS SALARIALES (Brutas) del Hogar (2010)'.

*----- MATCH - RASAL_B_m: RENTAS SALARIALES (Brutas) del Hogar para matching -----
Numeric RASAL_B_m (F10.2).
compute RASAL_B_m=RMCA_B+RNMCA_B.
execute.
Variable Labels RASAL_B_m 'RENTAS SALARIALES (Brutas) del Hogar para matching'.
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=RASAL_B_m ngrupos=10 pi=sí .

*----- IMCAP_B: Rentas del Capital Monetario (brutos) -----
Numeric IMCAP_B (F10.2).
compute IMCAP_B=0.
Missing Values P4_25 P4_16 P4_8_1 P4_8_2 P4_8_3 P4_36 P4_40 (-1, -2).
if (not(missing(P4_25))) IMCAP_B=IMCAP_B+P4_25 + 0000 .
if (not(missing(P4_16))) IMCAP_B=IMCAP_B+P4_16 + 00.
if (not(missing(P4_8_1))) IMCAP_B=IMCAP_B+P4_8_1.
if (not(missing(P4_8_2))) IMCAP_B=IMCAP_B+P4_8_2.
if (not(missing(P4_8_3))) IMCAP_B=IMCAP_B+P4_8_3.
if (not(missing(P4_36))) IMCAP_B=IMCAP_B+P4_36.
if (not(missing(P4_40))) IMCAP_B=IMCAP_B+P4_40.

```

```

execute .
Variable Labels IMCAP_B      'Rentas Brutas del Capital Monetario del hogar'.
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=IMCAP_B ngrupos=10 pi=sí .

Numeric IMCAP_BAA (F10.2).
compute IMCAP_BAA=0.
Missing Values P7_8a P7_8b P7_10 P7_12 P7_12a P7_14 (-1, -2).
if (not(missing(P7_8a))) IMCAP_BAA=IMCAP_BAA+P7_8a.
if (not(missing(P7_8b))) IMCAP_BAA=IMCAP_BAA+P7_8b      + 000.
if (not(missing(P7_10))) IMCAP_BAA=IMCAP_BAA+P7_10.
if (not(missing(P7_12))) IMCAP_BAA=IMCAP_BAA+P7_12      + 0000.
if (not(missing(P7_12a))) IMCAP_BAA=IMCAP_BAA+P7_12a.
if (not(missing(P7_14))) IMCAP_BAA=IMCAP_BAA+P7_14.
execute.
Variable Labels IMCAP_BAA 'Rentas Brutas del Capital (2010)'.

*----- IMPRO_B: Rentas de las propiedades o terrenos (Brutas) -----
Numeric IMPRO_B (F10.2).
compute IMPRO_B=0.
Missing Values p2_43_1 p2_43_2 p2_43_3 p2_43_4
                p2_48_1 p2_48_2 p2_48_3 p2_48_4      (-1, -2).
do repeat AAA  = p2_43_1 p2_43_2 p2_43_3 p2_43_4
              / BBB  = p2_48_1 p2_48_2 p2_48_3 p2_48_4.
              if (not(missing(AAA))) IMPRO_B=IMPRO_B+12*AAA.
              if (not(missing(BBB))) IMPRO_B=IMPRO_B-BBB.
end repeat.
execute .
Variable Labels IMPRO_B 'Rentas Brutas de las propiedades o terrenos'.
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=IMPRO_B ngrupos=10 pi=sí .

Numeric IMPRO_BAA (F10.2).
compute IMPRO_BAA=0.
Missing Values p7_2 p7_4a p7_4b p7_6a p7_6b      (-1, -2).
if (not(missing(P7_2))) IMPRO_BAA=IMPRO_BAA+P7_2.
if (not(missing(P7_4a))) IMPRO_BAA=IMPRO_BAA+P7_4a.
if (not(missing(P7_4b))) IMPRO_BAA=IMPRO_BAA+P7_4b.
if (not(missing(P7_6a))) IMPRO_BAA=IMPRO_BAA+P7_6a.
if (not(missing(P7_6b))) IMPRO_BAA=IMPRO_BAA+P7_6b.
execute.
Variable Labels IMPRO_BAA 'Rentas de la Propiedad (2010)'.

*----- RPROI_B: RENTAS DE LA PROPIEDAD (INGRESOS) (Brutos) -----
* Numeric RPROI_B (F10.2).
* compute RPROI_B=$SYSMIS.
* execute .
* Variable Labels RPROI_B      'RENTAS BRUTAS DE LA PROPIEDAD (INGRESOS) DEL HOGAR' .

*----- RPROG_B: RENTAS DE LA PROPIEDAD (GASTOS DEDUCIBLES) (Brutos) -----
* Numeric RPROG_B (F10.2).
* compute RPROG_B=$SYSMIS.
* execute .
* Variable Labels RPROG_B      'RENTAS BRUTAS DE LA PROPIEDAD (GASTOS DEDUCIBLES) DEL HOGAR'.

*----- RPRO_B: RENTAS DE LA PROPIEDAD (INGRESOS-GASTOS DEDUCIBLES) (Brutas) -----
Numeric RPRO_B (F10.2).
compute RPRO_B=IMCAP_B+IMPRO_B.
execute .
Variable Labels RPRO_B 'RENTAS BRUTAS DE LA PROPIEDAD (INGRESOS-GASTOS DEDUCIBLES) DEL HOGAR'.
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=RPRO_B ngrupos=10 pi=sí .

Numeric RPRO_BAA (F10.2).
compute RPRO_BAA=IMCAP_BAA+IMPRO_BAA.
execute .
Variable Labels RPRO_BAA 'RENTAS BRUTAS DE LA PROPIEDAD (INGRESOS-GASTOS DEDUCIBLES) DEL HOGAR (2010)'.

*----- ORNC_B: Otras Rentas no clasificadas (Brutas) -----
Numeric ORNC_B (F10.2).
compute ORNC_B=0.
execute .
Variable Labels ORNC_B      'Otras Rentas no clasificadas (Brutas) DEL HOGAR'.

Numeric ORNC_BAA (F10.2).
compute ORNC_BAA=0.
execute.
Variable Labels ORNC_BAA 'Otras Rentas no clasificadas (Brutas), inc Renta Extraordinaria (2010)'.

*----- RPRIH_B: SALDO DE RENTAS PRIMARIAS DEL HOGAR (sin IAVP) (Brutas) -----
Numeric RPRIH_B (F10.2).
compute RPRIH_B = RMIX_B+RASAL_B+RPRO_B      + ORNC_B.
execute.

```

```

Numeric RPRIH_BAA (F10.2).
compute RPRIH_BAA = RMIX_BAA+RASAL_BAA+RPRO_BAA +ORNC_BAA.
execute.
Variable Labels RPRIH_BAA 'SALDO DE RENTAS PRIMARIAS BRUTAS DEL HOGAR (2010) (sin IAVP)'.

*----- MATCH - RPRIH_B_m: SALDO DE RENTAS PRIMARIAS DEL HOGAR (sin IAVP) (Brutas) para matching -----
Numeric RPRIH_B_m (F10.2).
compute RPRIH_B_m = RMIX_B+RASAL_B_m+RPRO_B.
execute.
Variable Labels RPRIH_B_m 'SALDO DE RENTAS PRIMARIAS DEL HOGAR (sin IAVP) (Brutas) para matching'.
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=RPRIH_B_m ngrupos=10 pi=sí.

*----- PSDES_B: Prestaciones por desempleo (brutas) -----
Numeric PSDES_B (F10.2).
compute PSDES_B=0.
Missing Values p6_28b_1 p6_28b_2 p6_28b_3 p6_28b_4 p6_28b_5 p6_28b_6 p6_28b_7 p6_28b_8 p6_28b_9
p6_28c_1 p6_28c_2 p6_28c_3 p6_28c_4 p6_28c_5 p6_28c_6 p6_28c_7 p6_28c_8 p6_28c_9
p6_28c2_1 p6_28c2_2 p6_28c2_3 p6_28c2_4 p6_28c2_5 p6_28c2_6 p6_28c2_7 p6_28c2_8 p6_28c2_9 (-1, -2).
do repeat BBB = p6_28b_1 p6_28b_2 p6_28b_3 p6_28b_4 p6_28b_5 p6_28b_6 p6_28b_7 p6_28b_8 p6_28b_9
/ CCC = p6_28c_1 p6_28c_2 p6_28c_3 p6_28c_4 p6_28c_5 p6_28c_6 p6_28c_7 p6_28c_8 p6_28c_9
/ DDD = p6_28c2_1 p6_28c2_2 p6_28c2_3 p6_28c2_4 p6_28c2_5 p6_28c2_6 p6_28c2_7 p6_28c2_8 p6_28c2_9.
if (not(missing(BBB))) PSDES_B=PSDES_B+12*BBB.
if (not(missing(CCC))) PSDES_B=PSDES_B+12*CCC.
if (not(missing(DDD))) PSDES_B=PSDES_B+12*DDD.
end repeat.
Missing Values p6_28d_1 p6_28d_2 p6_28d_3 p6_28d_4 p6_28d_5 p6_28d_6 p6_28d_7 p6_28d_8 p6_28d_9 (-1, -2).
do repeat EEE = p6_28d_1 p6_28d_2 p6_28d_3 p6_28d_4 p6_28d_5 p6_28d_6 p6_28d_7 p6_28d_8 p6_28d_9.
if (not(missing(EEE))) PSDES_B=PSDES_B+12*EEE.
end repeat.
execute.
Variable Labels PSDES_B 'Prestaciones por desempleo (brutas) del Hogar'.
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=PSDES_B ngrupos=10 pi=sí.

Numeric PSDES_BAA (F10.2).
compute PSDES_BAA=0.
Missing Values p6_68_1 p6_68_2 p6_68_3 p6_68_4 p6_68_5 p6_68_6 p6_68_7 p6_68_8 p6_68_9 (-1, -2).
do repeat AAA = p6_68_1 p6_68_2 p6_68_3 p6_68_4 p6_68_5 p6_68_6 p6_68_7 p6_68_8 p6_68_9.
if (not(missing(AAA))) PSDES_BAA=PSDES_BAA+AAA.
end repeat.
execute.
Variable Labels PSDES_BAA 'Prestaciones sociales (Desempleo) del Hogar (2010)'.

*----- PSJUB_B: Prestaciones por jubilación / vejez (brutas) -----
Numeric PSJUB_B (F10.2).
compute PSJUB_B=0.
Missing Values p6_49_1_1 p6_49_2_1 p6_49_3_1 p6_49_4_1 p6_49_5_1 p6_49_6_1 p6_49_7_1 p6_49_8_1 p6_49_9_1
p6_49_1_2 p6_49_2_2 p6_49_3_2 p6_49_4_2 p6_49_5_2 p6_49_6_2 p6_49_7_2 p6_49_8_2 p6_49_9_2
p6_49_1_3 p6_49_2_3 p6_49_3_3 p6_49_4_3 p6_49_5_3 p6_49_6_3 p6_49_7_3 p6_49_8_3 p6_49_9_3
p6_49_1_4 p6_49_2_4 p6_49_3_4 p6_49_4_4 p6_49_5_4 p6_49_6_4 p6_49_7_4 p6_49_8_4 p6_49_9_4
p6_49b_1_1 p6_49b_2_1 p6_49b_3_1 p6_49b_4_1 p6_49b_5_1 p6_49b_6_1 p6_49b_7_1 p6_49b_8_1 p6_49b_9_1
p6_49b_1_2 p6_49b_2_2 p6_49b_3_2 p6_49b_4_2 p6_49b_5_2 p6_49b_6_2 p6_49b_7_2 p6_49b_8_2 p6_49b_9_2
p6_49b_1_3 p6_49b_2_3 p6_49b_3_3 p6_49b_4_3 p6_49b_5_3 p6_49b_6_3 p6_49b_7_3 p6_49b_8_3 p6_49b_9_3
p6_49b_1_4 p6_49b_2_4 p6_49b_3_4 p6_49b_4_4 p6_49b_5_4 p6_49b_6_4 p6_49b_7_4 p6_49b_8_4 p6_49b_9_4 (-1,
-2).
do repeat TTT = p6_47_1_1 p6_47_2_1 p6_47_3_1 p6_47_4_1 p6_47_5_1 p6_47_6_1 p6_47_7_1 p6_47_8_1 p6_47_9_1
p6_47_1_2 p6_47_2_2 p6_47_3_2 p6_47_4_2 p6_47_5_2 p6_47_6_2 p6_47_7_2 p6_47_8_2 p6_47_9_2
p6_47_1_3 p6_47_2_3 p6_47_3_3 p6_47_4_3 p6_47_5_3 p6_47_6_3 p6_47_7_3 p6_47_8_3 p6_47_9_3
p6_47_1_4 p6_47_2_4 p6_47_3_4 p6_47_4_4 p6_47_5_4 p6_47_6_4 p6_47_7_4 p6_47_8_4 p6_47_9_4
/ AAA = p6_49_1_1 p6_49_2_1 p6_49_3_1 p6_49_4_1 p6_49_5_1 p6_49_6_1 p6_49_7_1 p6_49_8_1 p6_49_9_1
p6_49_1_2 p6_49_2_2 p6_49_3_2 p6_49_4_2 p6_49_5_2 p6_49_6_2 p6_49_7_2 p6_49_8_2 p6_49_9_2
p6_49_1_3 p6_49_2_3 p6_49_3_3 p6_49_4_3 p6_49_5_3 p6_49_6_3 p6_49_7_3 p6_49_8_3 p6_49_9_3
p6_49_1_4 p6_49_2_4 p6_49_3_4 p6_49_4_4 p6_49_5_4 p6_49_6_4 p6_49_7_4 p6_49_8_4 p6_49_9_4
/ BBB = p6_49b_1_1 p6_49b_2_1 p6_49b_3_1 p6_49b_4_1 p6_49b_5_1 p6_49b_6_1 p6_49b_7_1 p6_49b_8_1 p6_49b_9_1
p6_49b_1_2 p6_49b_2_2 p6_49b_3_2 p6_49b_4_2 p6_49b_5_2 p6_49b_6_2 p6_49b_7_2 p6_49b_8_2 p6_49b_9_2
p6_49b_1_3 p6_49b_2_3 p6_49b_3_3 p6_49b_4_3 p6_49b_5_3 p6_49b_6_3 p6_49b_7_3 p6_49b_8_3 p6_49b_9_3
p6_49b_1_4 p6_49b_2_4 p6_49b_3_4 p6_49b_4_4 p6_49b_5_4 p6_49b_6_4 p6_49b_7_4 p6_49b_8_4 p6_49b_9_4.
do if (TTT=1).
if (not(missing(AAA))) PSJUB_B=PSJUB_B+12*AAA.
if (not(missing(BBB))) PSJUB_B=PSJUB_B+12*BBB.
end if.
end repeat.
execute.
Variable Labels PSJUB_B 'Prestaciones por jubilación / vejez (brutas) del Hogar'.
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=PSJUB_B ngrupos=10 pi=sí.

*----- PSSUP_B: Prestaciones por supervivencia (brutas) -----
Numeric PSSUP_B (F10.2).
compute PSSUP_B=0.
Missing Values P6_54 P6_53b (-1, -2).
if (not(missing(P6_54))) PSSUP_B=PSSUP_B+12*P6_54.
if (not(missing(P6_53b))) PSSUP_B=PSSUP_B+12*P6_53b.

```

```

execute .
Variable Labels PSSUP_B 'Prestaciones por supervivencia (brutas) del Hogar'.
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=PSSUP_B ngrupos=10 pi=sí .

*----- MATCH - PSJYS_B_m: Prestaciones por jubilación / vejez y supervivencia (brutas) para matching -----,
Numeric PSJYS_B_m (F10.2).
compute PSJYS_B_m = PSJUB_B+PSSUP_B .
execute.
Variable Labels PSJYS_B_m 'Prestaciones por jubilación / vejez y supervivencia (brutas) para matching'.
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=PSJYS_B_m ngrupos=10 pi=sí .

*----- PSENF_B: Prestaciones por enfermedad (brutas) -----,
Numeric PSENF_B (F10.2).
compute PSENF_B=0.
Missing Values P6_52 P6_51b (-1, -2).
if (not(missing(P6_52))) PSENF_B=PSENF_B+12*P6_52.
if (not(missing(P6_51b))) PSENF_B=PSENF_B-12*P6_51b.
execute .
Variable Labels PSENF_B 'Prestaciones por enfermedad (brutas) del Hogar'.
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=PSENF_B ngrupos=10 pi=sí .

*----- PSINV_B: Prestaciones por invalidez (brutas) -----,
Numeric PSINV_B (F10.2).
compute PSINV_B=0.
Missing Values p6_49_1_1 p6_49_2_1 p6_49_3_1 p6_49_4_1 p6_49_5_1 p6_49_6_1 p6_49_7_1 p6_49_8_1 p6_49_9_1
p6_49_1_2 p6_49_2_2 p6_49_3_2 p6_49_4_2 p6_49_5_2 p6_49_6_2 p6_49_7_2 p6_49_8_2 p6_49_9_2
p6_49_1_3 p6_49_2_3 p6_49_3_3 p6_49_4_3 p6_49_5_3 p6_49_6_3 p6_49_7_3 p6_49_8_3 p6_49_9_3
p6_49_1_4 p6_49_2_4 p6_49_3_4 p6_49_4_4 p6_49_5_4 p6_49_6_4 p6_49_7_4 p6_49_8_4 p6_49_9_4
p6_49b_1_1 p6_49b_2_1 p6_49b_3_1 p6_49b_4_1 p6_49b_5_1 p6_49b_6_1 p6_49b_7_1 p6_49b_8_1 p6_49b_9_1
p6_49b_1_2 p6_49b_2_2 p6_49b_3_2 p6_49b_4_2 p6_49b_5_2 p6_49b_6_2 p6_49b_7_2 p6_49b_8_2 p6_49b_9_2
p6_49b_1_3 p6_49b_2_3 p6_49b_3_3 p6_49b_4_3 p6_49b_5_3 p6_49b_6_3 p6_49b_7_3 p6_49b_8_3 p6_49b_9_3
p6_49b_1_4 p6_49b_2_4 p6_49b_3_4 p6_49b_4_4 p6_49b_5_4 p6_49b_6_4 p6_49b_7_4 p6_49b_8_4 p6_49b_9_4 (-1,
-2).
do repeat TTT = p6_47_1_1 p6_47_2_1 p6_47_3_1 p6_47_4_1 p6_47_5_1 p6_47_6_1 p6_47_7_1 p6_47_8_1 p6_47_9_1
p6_47_1_2 p6_47_2_2 p6_47_3_2 p6_47_4_2 p6_47_5_2 p6_47_6_2 p6_47_7_2 p6_47_8_2 p6_47_9_2
p6_47_1_3 p6_47_2_3 p6_47_3_3 p6_47_4_3 p6_47_5_3 p6_47_6_3 p6_47_7_3 p6_47_8_3 p6_47_9_3
p6_47_1_4 p6_47_2_4 p6_47_3_4 p6_47_4_4 p6_47_5_4 p6_47_6_4 p6_47_7_4 p6_47_8_4 p6_47_9_4
/ AAA = p6_49_1_1 p6_49_2_1 p6_49_3_1 p6_49_4_1 p6_49_5_1 p6_49_6_1 p6_49_7_1 p6_49_8_1 p6_49_9_1
p6_49_1_2 p6_49_2_2 p6_49_3_2 p6_49_4_2 p6_49_5_2 p6_49_6_2 p6_49_7_2 p6_49_8_2 p6_49_9_2
p6_49_1_3 p6_49_2_3 p6_49_3_3 p6_49_4_3 p6_49_5_3 p6_49_6_3 p6_49_7_3 p6_49_8_3 p6_49_9_3
p6_49_1_4 p6_49_2_4 p6_49_3_4 p6_49_4_4 p6_49_5_4 p6_49_6_4 p6_49_7_4 p6_49_8_4 p6_49_9_4
/ BBB = p6_49b_1_1 p6_49b_2_1 p6_49b_3_1 p6_49b_4_1 p6_49b_5_1 p6_49b_6_1 p6_49b_7_1 p6_49b_8_1 p6_49b_9_1
p6_49b_1_2 p6_49b_2_2 p6_49b_3_2 p6_49b_4_2 p6_49b_5_2 p6_49b_6_2 p6_49b_7_2 p6_49b_8_2 p6_49b_9_2
p6_49b_1_3 p6_49b_2_3 p6_49b_3_3 p6_49b_4_3 p6_49b_5_3 p6_49b_6_3 p6_49b_7_3 p6_49b_8_3 p6_49b_9_3
p6_49b_1_4 p6_49b_2_4 p6_49b_3_4 p6_49b_4_4 p6_49b_5_4 p6_49b_6_4 p6_49b_7_4 p6_49b_8_4 p6_49b_9_4.
do if (TTT=2).
if (not(missing(AAA))) PSINV_B=PSINV_B+12*AAA.
if (not(missing(BBB))) PSINV_B=PSINV_B-12*BBB.
end if.
end repeat.
execute .
Variable Labels PSINV_B 'Prestaciones por invalidez (brutas) del Hogar'.
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=PSINV_B ngrupos=10 pi=sí .

*----- MATCH - PSEEL_B_m: Prestaciones por enfermedad e invalidez (brutas) para matching -----,
Numeric PSEEL_B_m (F10.2).
compute PSEEL_B_m = PSENF_B+PSINV_B .
execute.
Variable Labels PSEEL_B_m 'Prestaciones por enfermedad e invalidez (brutas) para matching'.
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=PSEEL_B_m ngrupos=10 pi=sí .

*----- PSEST_B: Ayudas para estudios (brutas) -----,
Numeric PSEST_B (F10.2).
compute PSEST_B=0.
Missing Values P6_58 P6_57b (-1, -2).
if (not(missing(P6_58))) PSEST_B=PSEST_B+12*P6_58.
if (not(missing(P6_57b))) PSEST_B=PSEST_B-12*P6_57b.
execute.
Variable Labels PSEST_B 'Ayudas para estudios (brutas) del Hogar'.
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=PSEST_B ngrupos=10 pi=sí .

*----- MATCH - PSFYV_B_m: Ayudas por familia/hijos, incluida Ayudas para vivienda(brutas) para matching -----,
Numeric PSFYV_B_m (F10.2).
compute PSFYV_B_m=0.
Missing Values P6_56 P6_55b (-1, -2).
if (not(missing(P6_56))) PSFYV_B_m=PSFYV_B_m+12*P6_56.
if (not(missing(P6_55b))) PSFYV_B_m=PSFYV_B_m-12*P6_55b.
execute .
Variable Labels PSFYV_B_m 'Ayudas por familia/hijos, incluida Ayudas para vivienda(brutas) para matching'.
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=PSFYV_B_m ngrupos=10 pi=sí .

```

```

*----- PSONC_B: Otras ayudas sociales no clasificadas en otro apartado (bruta) -----,
***** >>>> (No se recogen en le encuesta) <<<< *****.
Numeric PSONC_B (F10.2).
compute PSONC_B=0.
execute .
Variable Labels PSONC_B      'Otras ayudas sociales al hogar no clasificadas en otro apartado (bruta)'.

*----- PSJVI_BAA: Prestaciones por jubilación / vejez e Incapacidad (brutas) (inc FPensiones) -----,
Numeric PSJVI_BAA (F10.2).
compute PSJVI_BAA=0.
Missing Values p6_74_1 p6_74_2 p6_74_3 p6_74_4 p6_74_5 p6_74_6 p6_74_7 p6_74_8 p6_74_9  (-1, -2).
do repeat AAA = p6_74_1 p6_74_2 p6_74_3 p6_74_4 p6_74_5 p6_74_6 p6_74_7 p6_74_8 p6_74_9 .
    if (not(missing(AAA))) PSJVI_BAA=PSJVI_BAA+AAA.
end repeat.
execute.
Variable Labels PSJVI_BAA 'Prestaciones Jubilación e Incapacidad (2010) (inc FPensiones)'.

*----- PlyANL_BAA: Indemnizaciones y Ayudas no laborales (bruta) -----,
Numeric PlyANL_BAA (F10.2).
compute PlyANL_BAA=0.
    if (leff='EFF08') !then
Missing Values P6_75b P6_75d  (-1, -2).
if (not(missing(P6_75b))) PlyANL_BAA=PlyANL_BAA+P6_75b.
if (not(missing(P6_75d))) PlyANL_BAA=PlyANL_BAA+P6_75d.
    !ifend
    if (leff='EFF11') !then
Missing Values P6_75b P6_75d1 P6_75d2 P6_75d3 P6_75d4  (-1, -2).
if (not(missing(P6_75b))) PlyANL_BAA=PlyANL_BAA+P6_75b.
if (not(missing(P6_75d1))) PlyANL_BAA=PlyANL_BAA+P6_75d1.
if (not(missing(P6_75d2))) PlyANL_BAA=PlyANL_BAA+P6_75d2.
if (not(missing(P6_75d3))) PlyANL_BAA=PlyANL_BAA+P6_75d3.
if (not(missing(P6_75d4))) PlyANL_BAA=PlyANL_BAA+P6_75d4.
    !ifend
execute.
Variable Labels PlyANL_BAA 'Indemnizaciones y Ayudas no laborales (2010)'.

*----- RPEPP_N: Rentas Netas percibidas de esquemas privados de pensiones (no SEEPROS) -----,
Numeric RPEPP_N (F10.2).
compute RPEPP_N=0.
Missing Values p6_49_1_1 p6_49_2_1 p6_49_3_1 p6_49_4_1 p6_49_5_1 p6_49_6_1 p6_49_7_1 p6_49_8_1 p6_49_9_1
    p6_49_1_2 p6_49_2_2 p6_49_3_2 p6_49_4_2 p6_49_5_2 p6_49_6_2 p6_49_7_2 p6_49_8_2 p6_49_9_2
    p6_49_1_3 p6_49_2_3 p6_49_3_3 p6_49_4_3 p6_49_5_3 p6_49_6_3 p6_49_7_3 p6_49_8_3 p6_49_9_3
    p6_49_1_4 p6_49_2_4 p6_49_3_4 p6_49_4_4 p6_49_5_4 p6_49_6_4 p6_49_7_4 p6_49_8_4 p6_49_9_4
    p6_49b_1_1 p6_49b_2_1 p6_49b_3_1 p6_49b_4_1 p6_49b_5_1 p6_49b_6_1 p6_49b_7_1 p6_49b_8_1 p6_49b_9_1
    p6_49b_1_2 p6_49b_2_2 p6_49b_3_2 p6_49b_4_2 p6_49b_5_2 p6_49b_6_2 p6_49b_7_2 p6_49b_8_2 p6_49b_9_2
    p6_49b_1_3 p6_49b_2_3 p6_49b_3_3 p6_49b_4_3 p6_49b_5_3 p6_49b_6_3 p6_49b_7_3 p6_49b_8_3 p6_49b_9_3
    p6_49b_1_4 p6_49b_2_4 p6_49b_3_4 p6_49b_4_4 p6_49b_5_4 p6_49b_6_4 p6_49b_7_4 p6_49b_8_4 p6_49b_9_4  (-1,
-2).
do repeat TTT = p6_47_1_1 p6_47_2_1 p6_47_3_1 p6_47_4_1 p6_47_5_1 p6_47_6_1 p6_47_7_1 p6_47_8_1 p6_47_9_1
    p6_47_1_2 p6_47_2_2 p6_47_3_2 p6_47_4_2 p6_47_5_2 p6_47_6_2 p6_47_7_2 p6_47_8_2 p6_47_9_2
    p6_47_1_3 p6_47_2_3 p6_47_3_3 p6_47_4_3 p6_47_5_3 p6_47_6_3 p6_47_7_3 p6_47_8_3 p6_47_9_3
    p6_47_1_4 p6_47_2_4 p6_47_3_4 p6_47_4_4 p6_47_5_4 p6_47_6_4 p6_47_7_4 p6_47_8_4 p6_47_9_4
    / AAA = p6_49_1_1 p6_49_2_1 p6_49_3_1 p6_49_4_1 p6_49_5_1 p6_49_6_1 p6_49_7_1 p6_49_8_1 p6_49_9_1
    p6_49_1_2 p6_49_2_2 p6_49_3_2 p6_49_4_2 p6_49_5_2 p6_49_6_2 p6_49_7_2 p6_49_8_2 p6_49_9_2
    p6_49_1_3 p6_49_2_3 p6_49_3_3 p6_49_4_3 p6_49_5_3 p6_49_6_3 p6_49_7_3 p6_49_8_3 p6_49_9_3
    p6_49_1_4 p6_49_2_4 p6_49_3_4 p6_49_4_4 p6_49_5_4 p6_49_6_4 p6_49_7_4 p6_49_8_4 p6_49_9_4
    / BBB = p6_49b_1_1 p6_49b_2_1 p6_49b_3_1 p6_49b_4_1 p6_49b_5_1 p6_49b_6_1 p6_49b_7_1 p6_49b_8_1 p6_49b_9_1
    p6_49b_1_2 p6_49b_2_2 p6_49b_3_2 p6_49b_4_2 p6_49b_5_2 p6_49b_6_2 p6_49b_7_2 p6_49b_8_2 p6_49b_9_2
    p6_49b_1_3 p6_49b_2_3 p6_49b_3_3 p6_49b_4_3 p6_49b_5_3 p6_49b_6_3 p6_49b_7_3 p6_49b_8_3 p6_49b_9_3
    p6_49b_1_4 p6_49b_2_4 p6_49b_3_4 p6_49b_4_4 p6_49b_5_4 p6_49b_6_4 p6_49b_7_4 p6_49b_8_4 p6_49b_9_4.
do if (TTT=3).
    if (not(missing(AAA))) PSEEL_B=PSEEL_B+12*AAA.
    if (not(missing(BBB))) PSEEL_B=PSEEL_B+12*BBB.
end if.
end repeat.
execute .
Variable Labels RPEPP_N      'Rentas netas percibidas de esquemas privados de pensiones (no SEEPROS)'.

*----- RPEPP_B: Rentas netas percibidas de esquemas privados de pensiones (no SEEPROS) -----,
Numeric RPEPP_B (F10.2).
compute RPEPP_B=RPEPP_N.
execute.
Variable Labels RPEPP_B      'Rentas netas percibidas de esquemas privados de pensiones (no SEEPROS)'.

!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=RPEPP_B ngrupos=10 pi=sí .

*----- PRESOC_B_m: PRESTACIONES SOCIALES (Brutas) para matching -----,
Numeric PRESOC_B_m (F10.2).
compute PRESOC_B_m = PSDES_B +PSJYS_B_m +PSEEL_B_m +PSEST_B +PSFYV_B_m.

```

```

execute.
Variable Labels PRESOC_B_m 'PRESTACIONES SOCIALES (Brutas) al hogar para matching'.
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=PRESOC_B_m ngrupos=10 pi=sí.

Numeric PRESOC_BAA (F10.2).
compute PRESOC_BAA = PSDES_BAA+PSJVI_BAA+PlyANL_BAA.
execute.
Variable Labels PRESOC_BAA 'PRESTACIONES SOCIALES (Brutas) al hogar (2010)'.

*----- TPRDOH_B: Transferencias periódicas monetarias brutas percibidas de otros hogares -----
Numeric TPRDOH_B (F10.2).
compute TPRDOH_B=0.
Missing Values p6_28a_1 p6_28a_2 p6_28a_3 p6_28a_4 p6_28a_5 p6_28a_6 p6_28a_7 p6_28a_8 p6_28a_9 (-1, -2).
do repeat AAA = p6_28a_1 p6_28a_2 p6_28a_3 p6_28a_4 p6_28a_5 p6_28a_6 p6_28a_7 p6_28a_8 p6_28a_9.
* / FFF = p6_28e_1 p6_28e_2 p6_28e_3 p6_28e_4 p6_28e_5 p6_28e_6 p6_28e_7 p6_28e_8 p6_28e_9.
* if (not(missing(AAA))) TPRDOH_B=TPRDOH_B+12*AAA.
* if (not(missing(FFF))) TPRDOH_B=TPRDOH_B+12*FFF + 000.
end repeat.
Missing Values P6_60 P6_59b (-1, -2).
if (not(missing(P6_60))) TPRDOH_B=TPRDOH_B+12*P6_60.
if (not(missing(P6_59b))) TPRDOH_B=TPRDOH_B-12*P6_59b.
execute.
Variable Labels TPRDOH_B 'Transferencias periódicas monetarias brutas percibidas de otros hogares'.
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=TPRDOH_B ngrupos=10 pi=sí.

Numeric TPRDOH_BAA (F10.2).
compute TPRDOH_BAA=0.
execute.
Variable Labels TPRDOH_BAA 'Transferencias periódicas monetarias brutas percibidas de otros hogares (2010)'.

*----- TCDOH_B: Transfer. Corrientes Percibidas de otros hogares (Brutas) -----
Numeric TCDOH_B (F10.2).
compute TCDOH_B = TPRDOH_B.
execute.
Variable Labels TCDOH_B 'Transfer. Corrientes Percibidas de otros hogares (Brutas) (Brutas) por el hogar'.

*----- TPCER_B_m: TRANSFERENCIAS CORRIENTES PERCIBIDAS (Brutas) para matching -----
Numeric TPCER_B_m (F10.2).
compute TPCER_B_m = PRESOC_B_m + TCDOH_B.
execute.
Variable Labels TPCER_B_m 'TRANSFERENCIAS CORRIENTES PERCIBIDAS (Brutas) por el hogar para matching'.
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=TPCER_B_m ngrupos=10 pi=sí.

Numeric TPCER_BAA (F10.2).
compute TPCER_BAA = PRESOC_BAA+TPRDOH_BAA.
execute.
Variable Labels TPCER_BAA 'TRANSFERENCIAS CORRIENTES PERCIBIDAS (Brutas) por el hogar (2010)'.

*----- REXT_B: Renta Extraordinaria (Bruta) del Hogar -----
Numeric REXT_B (F10.2).
compute REXT_B=0.
Missing Values P6_60f (-1, -2).
if (not(missing(P6_60f)) and not (P6_60es1=2 or P6_60es2=2)) REXT_B=REXT_B+4*P6_60f.
execute.
Variable Labels REXT_B 'Renta Extraordinaria (Bruta) del Hogar'.

Numeric REXT_BAA (F10.2).
compute REXT_BAA=0.
Missing Values p6_70_1 p6_70_2 p6_70_3 p6_70_4 p6_70_5 p6_70_6 p6_70_7 p6_70_8 p6_70_9 (-1, -2).
do repeat BBB = p6_70_1 p6_70_2 p6_70_3 p6_70_4 p6_70_5 p6_70_6 p6_70_7 p6_70_8 p6_70_9.
* if (not(missing(BBB))) REXT_BAA=REXT_BAA+BBB.
end repeat.
Missing Values P6_75f P9_14 P9_20 (-1, -2).
if (not(missing(P6_75f))) PlyANL_BAA=PlyANL_BAA+P6_75f + 000.
if (not(missing(P9_14)) and P9_13=2010) REXT_BAA=REXT_BAA+P9_14.
if (not(missing(P9_20)) and P9_19=2010) REXT_BAA=REXT_BAA+P9_20.
execute.
Variable Labels REXT_BAA 'Renta Extraordinaria (Bruta) del Hogar (2010)'.

*----- RT_B: RENTA TOTAL BRUTA DEL HOGAR (sin IAVP) -----
Numeric RT_B (F10.2).
compute RT_B = RPRIH_B_m + TPCER_B_m + REXT_B.
execute.
Variable Labels RT_B 'RENDA TOTAL BRUTA DEL HOGAR (sin IAVP)'.
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=RT_B ngrupos=10 pi=sí.

Numeric RT_BAA (F10.2).
compute RT_BAA = RPRIH_BAA + TPCER_BAA + REXT_BAA.
execute.
Variable Labels RT_BAA 'RENDA TOTAL BRUTA DEL HOGAR (sin IAVP) (2010)'.

```

```

*----- MATCH - INTOTALES_m: Total de Ingresos Netos del Hogar (No disponible) -----.

*----- MATCH - INTOTALES_m: Total de Ingresos Netos del Hogar (No disponible) -----.

*----- MATCH - LINTOTALES_m: Ingresos Netos del Hogar (No disponible) -----.

* ---- MATCH -FP_ING_m: Fuente Principal de ingresos del Hogar -----.
compute FP_ing_m=$SYSMIS.
compute #maxima=-1.e20.
compute #otrassub= PSEEL_B_m +PSEST_B +PSFYV_B_m .
compute #otras=TPRDOH_B+ REXT_B.
do repeat k=1      2      3      4      5      6      7
      / v= RMIX_B   RASAL_B_m PSJYS_B_m PSDES_B   #otrassub   RPRO_B   #otras.
do if (v>#maxima).
  compute FP_ing_m=k.
  compute #maxima=v.
end if.
end repeat.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=FP_ing_m
/TARGET VARIABLES=FP_ing_m
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

RECODE RMIX_B (-9=SYSMIS) (MISSING=SYSMIS) (0=0) (ELSE=1) INTO I_CPRO.
RECODE RASAL_B_m (-9=SYSMIS) (MISSING=SYSMIS) (0=0) (ELSE=1) INTO I_CAJE.
RECODE PSJYS_B_m (-9=SYSMIS) (MISSING=SYSMIS) (0=0) (ELSE=1) INTO I_PENS.
RECODE PSDES_B (-9=SYSMIS) (MISSING=SYSMIS) (0=0) (ELSE=1) INTO I_DES.
COMPUTE OTRSUB = PSEEL_B_m +PSEST_B +PSFYV_B_m.
RECODE OTRSUB (-9=SYSMIS) (MISSING=SYSMIS) (0=0) (ELSE=1) INTO I_OSPS.
RECODE RPRO_B (-9=SYSMIS) (MISSING=SYSMIS) (0=0) (ELSE=1) INTO I_RCYP.
COMPUTE OTROIN = TPRDOH_B+ REXT_B.
RECODE OTROIN (-9=SYSMIS) (MISSING=SYSMIS) (0=0) (ELSE=1) INTO I_OTRI.
EXECUTE.
VARIABLE LABELS
  I_DES 'El hogar tiene Ingresos por Desempleo'
  I_OSPS 'El hogar tiene Ingresos por Otras Subvenciones y PS'
  I_RCYP 'El hogar tiene Ingresos por Rentas del Capital y la Propiedad'
  I_OTRI 'El hogar tiene Ingresos por Otros Ingresos'
  I_CAJE 'El hogar tiene Ingresos por Cuenta Ajena'
  I_CPRO 'El hogar tiene Ingresos por Cuenta Propia'
  I_PENS 'El hogar tiene Ingresos por Pensiones C y NC'.
VALUE LABELS I_CPRO I_CAJE I_PENS I_DES I_OSPS I_RCYP I_OTRI
1 'Sí'
0 'No'.

===== COMPROBACIONES de las variables de Renta.
COMPUTE diff_renthogAA= RT_BAA-renthog.
execute.

SORT CASES BY diff_renthogAA(D).
DESCRIPTIVES VARIABLES=diff_renthogAA /STATISTICS=MEAN STDDEV MIN MAX.

WEIGHT BY w.
DESCRIPTIVES VARIABLES=nmh RMIX_BAA RASAL_BAA RPRO_BAA ORNC_BAA RPRIH_BAA PRESOC_BAA TPRDOH_BAA
TCPER_BAA REXT_BAA RT_BAA renthog
/STATISTICS=MEAN SUM STDDEV MIN MAX.
WEIGHT OFF.
delete variables diff_renthogAA.

=====.
** Variables sobre Gastos del Hogar .
=====.
*----- G_Consumo : Gasto Total de Consumo -----.
Numeric G_Consumo (F10.2).
compute G_Consumo=0.
Missing Values P9_1 (-1, -2).
if (not(missing(P9_1))) G_Consumo=P9_1*12.
execute.
Variable labels G_Consumo 'Gasto Anual Total de Consumo'.

*----- GASTOH : Gasto Total del Hogar -----.
Numeric GASTOH (F10.2).
compute GASTOH = G_Consumo.
execute.
Variable labels GASTOH 'Gasto Anual Total del Hogar'.
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=GASTOH ngrupos=10 pi=sí .

*----- G_alim: Gasto Total en Alimentación -----.
Numeric G_alim (F10.2).
compute G_alim=0.

```

```

Missing Values P9_2 (-1, -2).
if (not(missing(P9_2))) G_alim=P9_2*52*(P9_2b=1)+P9_2*12*(P9_2b=2).
execute.
Variable labels G_alim 'Gasto Anual en Alimentación'.

*----- GCF_01: Gasto Consumo Final Grupo 01 (Alimentación) -----,
Numeric GCF_01 (F10.2).
compute GCF_01=G_alim.
execute.
Variable labels GCF_01 'Gasto Consumo Final en Alimentación'.
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=GCF_01 ngrupos=10 pi=sí.

*----- G_OBDURA : Gasto en Otros Bienes duraderos -----,
Numeric G_OBDURA (F10.2).
compute G_OBDURA=$SYSMIS.
if (P2_69=2) G_OBDURA=0.
Missing Values P2_70 (-1, -2).
if (P2_69=1 and not(missing(P2_70))) G_OBDURA=P2_70.
execute.
Variable labels G_OBDURA 'Gasto Anual en Otros Bienes duraderos'.

*----- G_TCAOH: Transferencias a otros hogares -----,
Numeric G_TCAOH (F10.2).
compute G_TCAOH=$SYSMIS.
if (P9_3=2) G_TCAOH=0.
Missing Values P9_4 (-1, -2).
if (P9_3=1 and not(missing(P9_4))) G_TCAOH=G_TCAOH+P9_4*12.
execute.
Variable labels G_TCAOH 'Transferencias a otros hogares durante el año'.

*----- G_PPEN: Aportacion a Planes de Pensiones -----,
Numeric G_PPEN (F10.2).
compute G_PPEN=0.
Missing Values p5_6_1 p5_6_2 p5_6_3 p5_6_4 p5_6_5 p5_6_6 p5_6_7 p5_6_8 p5_6_9 p5_6_10 (-1, -2).
do repeat AAA=p5_6_1 p5_6_2 p5_6_3 p5_6_4 p5_6_5 p5_6_6 p5_6_7 p5_6_8 p5_6_9 p5_6_10 .
if (not(missing(AAA))) G_PPEN =G_PPEN + AAA .
end repeat.
execute.
Variable labels G_PPEN 'Aportacion Anual a Planes de Pensiones'.

*----- G_SEGVIDA: Aportacion Anual a Seguros de Vida -----,
Numeric G_SEGVIDA (F10.2).
compute G_SEGVIDA=0.
Missing Values p5_17a_1 p5_17a_2 p5_17a_3 p5_17a_4 p5_17a_5 p5_17a_6 (-1, -2).
do repeat AAA=p5_17a_1 p5_17a_2 p5_17a_3 p5_17a_4 p5_17a_5 p5_17a_6 .
if (not(missing(AAA))) G_SEGVIDA = G_SEGVIDA + AAA .
end repeat.
Missing Values p5_20b P5_23 (-1, -2).
if (not(missing(p5_20b))) G_SEGVIDA = G_SEGVIDA +p5_20b .
if (not(missing(P5_23))) G_SEGVIDA = G_SEGVIDA + P5_23 .
execute.
Variable labels G_SEGVIDA 'Aportacion Anual a Seguros de Vida'.

*----- PAGOSH: Total de Pagos anual del Hogar -----,
compute PAGOSH= pagodeuda+alim+nodur+gimpvehic+G_OBDURA+G_TCAOH+G_PPEN+G_SEGVIDA.
execute.
Variable Labels PAGOSH 'Total de Pagos anual del Hogar'.

*=====,
* Identificacion de las personas mas caracteristicas del Hogar .
*=====,

*----- cod_SP: Identificación del Sustentador Principal (SP) -----,
*----- cod_2A: Identificación del segundo Adulto (2A) -----,

Vector #aportac(9) .
loop #i=1 to 9.
compute #aportac(#i)=0.
end loop.

*----- RMCP_B : Renta Monetaria por Cuenta Propia (bruta) -----,
do repeat Aportac=#aportac1 #aportac2 #aportac3 #aportac4 #aportac5 #aportac6 #aportac7 #aportac8 #aportac9
/AAA1 = p6_102_1_1 p6_102_2_1 p6_102_3_1 p6_102_4_1 p6_102_5_1 p6_102_6_1 p6_102_7_1 p6_102_8_1
p6_102_9_1
/AAA2 = p6_102_1_2 p6_102_2_2 p6_102_3_2 p6_102_4_2 p6_102_5_2 p6_102_6_2 p6_102_7_2 p6_102_8_2
p6_102_9_2
/AAA3 = p6_102_1_3 p6_102_2_3 p6_102_3_3 p6_102_4_3 p6_102_5_3 p6_102_6_3 p6_102_7_3 p6_102_8_3
p6_102_9_3
/BBB1 = p6_104_1_1 p6_104_2_1 p6_104_3_1 p6_104_4_1 p6_104_5_1 p6_104_6_1 p6_104_7_1 p6_104_8_1 p6_104_9_1

```



```

/BBB2 = p6_104_1_2 p6_104_2_2 p6_104_3_2 p6_104_4_2 p6_104_5_2 p6_104_6_2 p6_104_7_2 p6_104_8_2 p6_104_9_2
/BBB3 = p6_104_1_3 p6_104_2_3 p6_104_3_3 p6_104_4_3 p6_104_5_3 p6_104_6_3 p6_104_7_3 p6_104_8_3
p6_104_9_3.
if (not(missing(AAA1))) Aportac=Aportac+12*AAA1.
if (not(missing(AAA2))) Aportac=Aportac+12*AAA2.
if (not(missing(AAA3))) Aportac=Aportac+12*AAA3.
if (not(missing(BBB1))) Aportac=Aportac+BBB1.
if (not(missing(BBB2))) Aportac=Aportac+BBB2.
if (not(missing(BBB3))) Aportac=Aportac+BBB3.
end repeat.
do repeat AAA = p4_112_1 p4_112_2 p4_112_3 p4_112_4 p4_112_5 p4_112_6
/ BBB = p4_112b_1 p4_112b_2 p4_112b_3 p4_112b_4 p4_112b_5 p4_112b_6
/ CCC = p4_103s1_1 p4_103s1_2 p4_103s1_3 p4_103s1_4 p4_103s1_5 p4_103s1_6 .
loop #1=1 to 9.
do if (#1 = CCC).
if (not(missing(AAA))) #aportac( #1 ) = #aportac ( #1 ) + AAA.
if (not(missing(BBB))) #aportac( #1 ) = #aportac ( #1 ) - BBB.
end if.
end loop.
end repeat.

*----- RNMCP_B: Renta No Monetaria por Cuenta Propia (se recogen neta) -----.
loop #1=1 to 9.
do if (#1 = p4_103s1_1).
if (not(missing(P6_60f)) and (P6_60es1=2 or P6_60es2=0)) #aportac(#1)=#aportac(#1)+4*P6_60f.
end if.
end loop.

*----- RMCA_B: Renta monetaria o cuasimonetaria de asalariados (bruta) -----.
compute #ind=0.
do repeat Aportac=#aportac1 #aportac2 #aportac3 #aportac4 #aportac5 #aportac6 #aportac7 #aportac8 #aportac9
/AAA1 = p6_14_1_1 p6_14_2_1 p6_14_3_1 p6_14_4_1 p6_14_5_1 p6_14_6_1 p6_14_7_1 p6_14_8_1 p6_14_9_1
/AAA2 = p6_14_1_2 p6_14_2_2 p6_14_3_2 p6_14_4_2 p6_14_5_2 p6_14_6_2 p6_14_7_2 p6_14_8_2 p6_14_9_2
/AAA3 = p6_14_1_3 p6_14_2_3 p6_14_3_3 p6_14_4_3 p6_14_5_3 p6_14_6_3 p6_14_7_3 p6_14_8_3 p6_14_9_3
/BBB1 = p6_14b_1_1 p6_14b_2_1 p6_14b_3_1 p6_14b_4_1 p6_14b_5_1 p6_14b_6_1 p6_14b_7_1 p6_14b_8_1 p6_14b_9_1
/BBB2 = p6_14b_1_2 p6_14b_2_2 p6_14b_3_2 p6_14b_4_2 p6_14b_5_2 p6_14b_6_2 p6_14b_7_2 p6_14b_8_2 p6_14b_9_2
/BBB3 = p6_14b_1_3 p6_14b_2_3 p6_14b_3_3 p6_14b_4_3 p6_14b_5_3 p6_14b_6_3 p6_14b_7_3 p6_14b_8_3 p6_14b_9_3.
do if (not(missing(AAA1)) and #ind eq 0).
if (not(missing(P6_60c))) Aportac = Aportac+12*P6_60c.
if (not(missing(P6_60b))) Aportac = Aportac-12*P6_60b.
compute #ind = 1.
end if.
if (not(missing(AAA1))) Aportac=Aportac+12*AAA1.
if (not(missing(AAA2))) Aportac=Aportac+12*AAA2.
if (not(missing(AAA3))) Aportac=Aportac+12*AAA3.
if (not(missing(BBB1))) Aportac=Aportac-12*BBB1.
if (not(missing(BBB2))) Aportac=Aportac-12*BBB2.
if (not(missing(BBB3))) Aportac=Aportac-12*BBB3.
end repeat.

*----- RNMCA_B: Renta bruta no monetaria de asalariados -----.
do repeat Aportac=#aportac1 #aportac2 #aportac3 #aportac4 #aportac5 #aportac6 #aportac7 #aportac8 #aportac9
/AAA1 = p6_16_1_1 p6_16_2_1 p6_16_3_1 p6_16_4_1 p6_16_5_1 p6_16_6_1 p6_16_7_1 p6_16_8_1 p6_16_9_1
/AAA2 = p6_16_1_2 p6_16_2_2 p6_16_3_2 p6_16_4_2 p6_16_5_2 p6_16_6_2 p6_16_7_2 p6_16_8_2 p6_16_9_2
/AAA3 = p6_16_1_3 p6_16_2_3 p6_16_3_3 p6_16_4_3 p6_16_5_3 p6_16_6_3 p6_16_7_3 p6_16_8_3 p6_16_9_3.
if (not(missing(AAA1))) Aportac=Aportac+12*AAA1.
if (not(missing(AAA2))) Aportac=Aportac+12*AAA2.
if (not(missing(AAA3))) Aportac=Aportac+12*AAA3.
end repeat.

*----- IMCAP_B: Rentas del Capital Monetario (brutas) -----.
if (not(missing(P4_25))) #aportac1= #aportac1+P4_25 + 0000 .
if (not(missing(P4_16))) #aportac1= #aportac1+P4_16 + 00.
if (not(missing(P4_8_1))) #aportac1= #aportac1+P4_8_1.
if (not(missing(P4_8_2))) #aportac1= #aportac1+P4_8_2.
if (not(missing(P4_8_3))) #aportac1= #aportac1+P4_8_3.
if (not(missing(P4_36))) #aportac1= #aportac1+P4_36.
if (not(missing(P4_40))) #aportac1= #aportac1+P4_40.

*----- IMPRO_B: Rentas de las propiedades o terrenos (Brutas) -----.
do repeat AAA = p2_43_1 p2_43_2 p2_43_3 p2_43_4
/ BBB = p2_48_1 p2_48_2 p2_48_3 p2_48_4.
if (not(missing(AAA))) #aportac1= #aportac1+12*AAA.
if (not(missing(BBB))) #aportac1= #aportac1 - BBB.
end repeat.

*----- PSDES_B: Prestaciones por desempleo (brutas) -----.
do repeat Aportac=#aportac1 #aportac2 #aportac3 #aportac4 #aportac5 #aportac6 #aportac7 #aportac8 #aportac9
/BBB = p6_28b_1 p6_28b_2 p6_28b_3 p6_28b_4 p6_28b_5 p6_28b_6 p6_28b_7 p6_28b_8 p6_28b_9
/CCC = p6_28c_1 p6_28c_2 p6_28c_3 p6_28c_4 p6_28c_5 p6_28c_6 p6_28c_7 p6_28c_8 p6_28c_9
/DDD = p6_28c2_1 p6_28c2_2 p6_28c2_3 p6_28c2_4 p6_28c2_5 p6_28c2_6 p6_28c2_7 p6_28c2_8 p6_28c2_9 .

```

```

        if (not(missing(BBB))) Aportac=Aportac+12*BBB.
        if (not(missing(CCC))) Aportac=Aportac+12*CCC.
        if (not(missing(DDD))) Aportac=Aportac-12*DDD.
    end repeat.
do repeat Aportac=#aportac1 #aportac2 #aportac3 #aportac4 #aportac5 #aportac6 #aportac7 #aportac8 #aportac9
    /EEE = p6_28d_1 p6_28d_2 p6_28d_3 p6_28d_4 p6_28d_5 p6_28d_6 p6_28d_7 p6_28d_8 p6_28d_9.
    if (not(missing(EEE))) Aportac=Aportac+12*EEE.
end repeat.

*----- PSJUB_B: Prestaciones por jubilación / vejez (brutas) -----
do repeat Aportac=#aportac1 #aportac2 #aportac3 #aportac4 #aportac5 #aportac6 #aportac7 #aportac8 #aportac9
    /TTT1 = p6_47_1_1 p6_47_2_1 p6_47_3_1 p6_47_4_1 p6_47_5_1 p6_47_6_1 p6_47_7_1 p6_47_8_1 p6_47_9_1
    /TTT2 = p6_47_1_2 p6_47_2_2 p6_47_3_2 p6_47_4_2 p6_47_5_2 p6_47_6_2 p6_47_7_2 p6_47_8_2 p6_47_9_2
    /TTT3 = p6_47_1_3 p6_47_2_3 p6_47_3_3 p6_47_4_3 p6_47_5_3 p6_47_6_3 p6_47_7_3 p6_47_8_3 p6_47_9_3
    /TTT4 = p6_47_1_4 p6_47_2_4 p6_47_3_4 p6_47_4_4 p6_47_5_4 p6_47_6_4 p6_47_7_4 p6_47_8_4 p6_47_9_4
    /AAA1 = p6_49_1_1 p6_49_2_1 p6_49_3_1 p6_49_4_1 p6_49_5_1 p6_49_6_1 p6_49_7_1 p6_49_8_1 p6_49_9_1
    /AAA2 = p6_49_1_2 p6_49_2_2 p6_49_3_2 p6_49_4_2 p6_49_5_2 p6_49_6_2 p6_49_7_2 p6_49_8_2 p6_49_9_2
    /AAA3 = p6_49_1_3 p6_49_2_3 p6_49_3_3 p6_49_4_3 p6_49_5_3 p6_49_6_3 p6_49_7_3 p6_49_8_3 p6_49_9_3
    /AAA4 = p6_49_1_4 p6_49_2_4 p6_49_3_4 p6_49_4_4 p6_49_5_4 p6_49_6_4 p6_49_7_4 p6_49_8_4 p6_49_9_4
    /BBB1 = p6_49b_1_1 p6_49b_2_1 p6_49b_3_1 p6_49b_4_1 p6_49b_5_1 p6_49b_6_1 p6_49b_7_1 p6_49b_8_1 p6_49b_9_1
    /BBB2 = p6_49b_1_2 p6_49b_2_2 p6_49b_3_2 p6_49b_4_2 p6_49b_5_2 p6_49b_6_2 p6_49b_7_2 p6_49b_8_2 p6_49b_9_2
    /BBB3 = p6_49b_1_3 p6_49b_2_3 p6_49b_3_3 p6_49b_4_3 p6_49b_5_3 p6_49b_6_3 p6_49b_7_3 p6_49b_8_3 p6_49b_9_3
    /BBB4 = p6_49b_1_4 p6_49b_2_4 p6_49b_3_4 p6_49b_4_4 p6_49b_5_4 p6_49b_6_4 p6_49b_7_4 p6_49b_8_4 p6_49b_9_4.
    if ((TTT1=1) and not(missing(AAA1))) Aportac=Aportac+12*AAA1.
    if ((TTT2=1) and not(missing(AAA2))) Aportac=Aportac+12*AAA2.
    if ((TTT3=1) and not(missing(AAA3))) Aportac=Aportac+12*AAA3.
    if ((TTT4=1) and not(missing(AAA4))) Aportac=Aportac+12*AAA4.
    if ((TTT1=1) and not(missing(BBB1))) Aportac=Aportac-12*BBB1.
    if ((TTT2=1) and not(missing(BBB2))) Aportac=Aportac-12*BBB2.
    if ((TTT3=1) and not(missing(BBB3))) Aportac=Aportac-12*BBB3.
    if ((TTT4=1) and not(missing(BBB4))) Aportac=Aportac-12*BBB4.
end repeat.

*----- PSSUP_B: Prestaciones por supervivencia (brutas) -----
if (not(missing(P6_54))) #aportac1=#aportac1+12*P6_54.
if (not(missing(P6_53b))) #aportac1=#aportac1-12*P6_53b.

*----- PSENF_B: Prestaciones por enfermedad (brutas) -----
if (not(missing(P6_52))) #aportac1=#aportac1+12*P6_52.
if (not(missing(P6_51b))) #aportac1=#aportac1-12*P6_51b.

*----- PSINV_B: Prestaciones por invalidez (brutas) -----
do repeat Aportac=#aportac1 #aportac2 #aportac3 #aportac4 #aportac5 #aportac6 #aportac7 #aportac8 #aportac9
    /TTT1 = p6_47_1_1 p6_47_2_1 p6_47_3_1 p6_47_4_1 p6_47_5_1 p6_47_6_1 p6_47_7_1 p6_47_8_1 p6_47_9_1
    /TTT2 = p6_47_1_2 p6_47_2_2 p6_47_3_2 p6_47_4_2 p6_47_5_2 p6_47_6_2 p6_47_7_2 p6_47_8_2 p6_47_9_2
    /TTT3 = p6_47_1_3 p6_47_2_3 p6_47_3_3 p6_47_4_3 p6_47_5_3 p6_47_6_3 p6_47_7_3 p6_47_8_3 p6_47_9_3
    /TTT4 = p6_47_1_4 p6_47_2_4 p6_47_3_4 p6_47_4_4 p6_47_5_4 p6_47_6_4 p6_47_7_4 p6_47_8_4 p6_47_9_4
    /AAA1 = p6_49_1_1 p6_49_2_1 p6_49_3_1 p6_49_4_1 p6_49_5_1 p6_49_6_1 p6_49_7_1 p6_49_8_1 p6_49_9_1
    /AAA2 = p6_49_1_2 p6_49_2_2 p6_49_3_2 p6_49_4_2 p6_49_5_2 p6_49_6_2 p6_49_7_2 p6_49_8_2 p6_49_9_2
    /AAA3 = p6_49_1_3 p6_49_2_3 p6_49_3_3 p6_49_4_3 p6_49_5_3 p6_49_6_3 p6_49_7_3 p6_49_8_3 p6_49_9_3
    /AAA4 = p6_49_1_4 p6_49_2_4 p6_49_3_4 p6_49_4_4 p6_49_5_4 p6_49_6_4 p6_49_7_4 p6_49_8_4 p6_49_9_4
    /BBB1 = p6_49b_1_1 p6_49b_2_1 p6_49b_3_1 p6_49b_4_1 p6_49b_5_1 p6_49b_6_1 p6_49b_7_1 p6_49b_8_1 p6_49b_9_1
    /BBB2 = p6_49b_1_2 p6_49b_2_2 p6_49b_3_2 p6_49b_4_2 p6_49b_5_2 p6_49b_6_2 p6_49b_7_2 p6_49b_8_2 p6_49b_9_2
    /BBB3 = p6_49b_1_3 p6_49b_2_3 p6_49b_3_3 p6_49b_4_3 p6_49b_5_3 p6_49b_6_3 p6_49b_7_3 p6_49b_8_3 p6_49b_9_3
    /BBB4 = p6_49b_1_4 p6_49b_2_4 p6_49b_3_4 p6_49b_4_4 p6_49b_5_4 p6_49b_6_4 p6_49b_7_4 p6_49b_8_4 p6_49b_9_4.
    if ((TTT1=2) and not(missing(AAA1))) Aportac=Aportac+12*AAA1.
    if ((TTT2=2) and not(missing(AAA2))) Aportac=Aportac+12*AAA2.
    if ((TTT3=2) and not(missing(AAA3))) Aportac=Aportac+12*AAA3.
    if ((TTT4=2) and not(missing(AAA4))) Aportac=Aportac+12*AAA4.
    if ((TTT1=2) and not(missing(BBB1))) Aportac=Aportac-12*BBB1.
    if ((TTT2=2) and not(missing(BBB2))) Aportac=Aportac-12*BBB2.
    if ((TTT3=2) and not(missing(BBB3))) Aportac=Aportac-12*BBB3.
    if ((TTT4=2) and not(missing(BBB4))) Aportac=Aportac-12*BBB4.
end repeat.

*----- PSEST_B: Ayudas para estudios (brutas) -----
if (not(missing(P6_58))) #aportac1=#aportac1+12*P6_58.
if (not(missing(P6_57b))) #aportac1=#aportac1-12*P6_57b.

*----- MATCH - PSFYV_B_m: Ayudas por familia/hijos, incluida Ayudas para vivienda(brutas) para matching -----
--
if (not(missing(P6_56))) #aportac1=#aportac1+12*P6_56.
if (not(missing(P6_55b))) #aportac1=#aportac1-12*P6_55b.

*----- TPRDOH_B: Transferencias periódicas monetarias brutas percibidas de otros hogares -----
do repeat Aportac=#aportac1 #aportac2 #aportac3 #aportac4 #aportac5 #aportac6 #aportac7 #aportac8 #aportac9
    /AAA = p6_28a_1 p6_28a_2 p6_28a_3 p6_28a_4 p6_28a_5 p6_28a_6 p6_28a_7 p6_28a_8 p6_28a_9.
    if (not(missing(AAA))) Aportac=Aportac+12*AAA.
end repeat.
if (not(missing(P6_60))) #aportac1= #aportac1+12*P6_60.
if (not(missing(P6_59b))) #aportac1= #aportac1-12*P6_59b.

```

```

*----- RPEPP_N: Rentas Netas percibidas de esquemas privados de pensiones (no SEEPROS) -----
do repeat Aportac=#aportac1 #aportac2 #aportac3 #aportac4 #aportac5 #aportac6 #aportac7 #aportac8 #aportac9
  /TTT1 = p6_47_1_1 p6_47_2_1 p6_47_3_1 p6_47_4_1 p6_47_5_1 p6_47_6_1 p6_47_7_1 p6_47_8_1 p6_47_9_1
  /TTT2 = p6_47_1_2 p6_47_2_2 p6_47_3_2 p6_47_4_2 p6_47_5_2 p6_47_6_2 p6_47_7_2 p6_47_8_2 p6_47_9_2
  /TTT3 = p6_47_1_3 p6_47_2_3 p6_47_3_3 p6_47_4_3 p6_47_5_3 p6_47_6_3 p6_47_7_3 p6_47_8_3 p6_47_9_3
  /TTT4 = p6_47_1_4 p6_47_2_4 p6_47_3_4 p6_47_4_4 p6_47_5_4 p6_47_6_4 p6_47_7_4 p6_47_8_4 p6_47_9_4
  /AAA1 = p6_49_1_1 p6_49_2_1 p6_49_3_1 p6_49_4_1 p6_49_5_1 p6_49_6_1 p6_49_7_1 p6_49_8_1 p6_49_9_1
  /AAA2 = p6_49_1_2 p6_49_2_2 p6_49_3_2 p6_49_4_2 p6_49_5_2 p6_49_6_2 p6_49_7_2 p6_49_8_2 p6_49_9_2
  /AAA3 = p6_49_1_3 p6_49_2_3 p6_49_3_3 p6_49_4_3 p6_49_5_3 p6_49_6_3 p6_49_7_3 p6_49_8_3 p6_49_9_3
  /AAA4 = p6_49_1_4 p6_49_2_4 p6_49_3_4 p6_49_4_4 p6_49_5_4 p6_49_6_4 p6_49_7_4 p6_49_8_4 p6_49_9_4
  /BBB1 = p6_49b_1_1 p6_49b_2_1 p6_49b_3_1 p6_49b_4_1 p6_49b_5_1 p6_49b_6_1 p6_49b_7_1 p6_49b_8_1 p6_49b_9_1
  /BBB2 = p6_49b_1_2 p6_49b_2_2 p6_49b_3_2 p6_49b_4_2 p6_49b_5_2 p6_49b_6_2 p6_49b_7_2 p6_49b_8_2 p6_49b_9_2
  /BBB3 = p6_49b_1_3 p6_49b_2_3 p6_49b_3_3 p6_49b_4_3 p6_49b_5_3 p6_49b_6_3 p6_49b_7_3 p6_49b_8_3 p6_49b_9_3
  /BBB4 = p6_49b_1_4 p6_49b_2_4 p6_49b_3_4 p6_49b_4_4 p6_49b_5_4 p6_49b_6_4 p6_49b_7_4 p6_49b_8_4 p6_49b_9_4
  if ((TTT1=3) and not(missing(AAA1))) Aportac=Aportac+12*AAA1.
  if ((TTT2=3) and not(missing(AAA2))) Aportac=Aportac+12*AAA2.
  if ((TTT3=3) and not(missing(AAA3))) Aportac=Aportac+12*AAA3.
  if ((TTT4=3) and not(missing(AAA4))) Aportac=Aportac+12*AAA4.
  if ((TTT1=3) and not(missing(BBB1))) Aportac=Aportac-12*BBB1.
  if ((TTT2=3) and not(missing(BBB2))) Aportac=Aportac-12*BBB2.
  if ((TTT3=3) and not(missing(BBB3))) Aportac=Aportac-12*BBB3.
  if ((TTT4=3) and not(missing(BBB4))) Aportac=Aportac-12*BBB4.
end repeat.

*----- REXT_B: Renta Extraordinaria (Bruta) del Hogar -----
if (not(missing(P6_60f)) and not (P6_60es1=2 or P6_60es2=2)) #aportac1=#aportac1+4*P6_60f.

vector ingr(9).
loop #i=1 to 9.
  compute ingr(#i) = #aportac(#i).
end loop.

vector #edadp(9).
do repeat edadp = #edadp1 #edadp2 #edadp3 #edadp4 #edadp5 #edadp6 #edadp7 #edadp8 #edadp9
  / v = p1_2d_1 p1_2d_2 p1_2d_3 p1_2d_4 p1_2d_5 p1_2d_6 p1_2d_7 p1_2d_8 p1_2d_9 .
  compute edadp = v.
end repeat.

vector #sexop(9).
do repeat sexop = #sexop1 #sexop2 #sexop3 #sexop4 #sexop5 #sexop6 #sexop7 #sexop8 #sexop9
  / v = p1_1_1 p1_1_2 p1_1_3 p1_1_4 p1_1_5 p1_1_6 p1_1_7 p1_1_8 p1_1_9 .
  compute sexop = v.
end repeat.

vector #sitacp(9).
do repeat v1 = p6_1c1_1 p6_1c1_2 p6_1c1_3 p6_1c1_4 p6_1c1_5 p6_1c1_6 p6_1c1_7 p6_1c1_8 p6_1c1_9
  / v2 = p6_1c2_1 p6_1c2_2 p6_1c2_3 p6_1c2_4 p6_1c2_5 p6_1c2_6 p6_1c2_7 p6_1c2_8 p6_1c2_9
  / v3 = p6_1c3_1 p6_1c3_2 p6_1c3_3 p6_1c3_4 p6_1c3_5 p6_1c3_6 p6_1c3_7 p6_1c3_8 p6_1c3_9
  / v4 = p6_1c4_1 p6_1c4_2 p6_1c4_3 p6_1c4_4 p6_1c4_5 p6_1c4_6 p6_1c4_7 p6_1c4_8 p6_1c4_9
  / v5 = p6_1c5_1 p6_1c5_2 p6_1c5_3 p6_1c5_4 p6_1c5_5 p6_1c5_6 p6_1c5_7 p6_1c5_8 p6_1c5_9
  / v6 = p6_1c6_1 p6_1c6_2 p6_1c6_3 p6_1c6_4 p6_1c6_5 p6_1c6_6 p6_1c6_7 p6_1c6_8 p6_1c6_9
  / v7 = p6_1c7_1 p6_1c7_2 p6_1c7_3 p6_1c7_4 p6_1c7_5 p6_1c7_6 p6_1c7_7 p6_1c7_8 p6_1c7_9
  / v8 = p6_1c8_1 p6_1c8_2 p6_1c8_3 p6_1c8_4 p6_1c8_5 p6_1c8_6 p6_1c8_7 p6_1c8_8 p6_1c8_9
  / sitacp = #sitacp1 #sitacp2 #sitacp3 #sitacp4 #sitacp5 #sitacp6 #sitacp7 #sitacp8 #sitacp9 .
  compute sitacp=$sysmis.
  do if (v1=1 or v2=1).
    compute sitacp=1 /* 1- ocupado */ .
  else if (v4=1 or v5=1).
    compute sitacp=4 /* 4- retirado */ .
  else if (v6=1).
    compute sitacp=3 /* 3- estudiante */ .
  else if (v3=1).
    compute sitacp=2 /* 2- desempleado */ .
  else if (v7=1).
    compute sitacp=5 /* 5 - labores del Hogar */ .
  else if (v8=1).
    compute sitacp=6 /* 6 - otra inactividad */ .
  end if.
end repeat.

compute #n_adulnodep=0.
vector #adulnodep(9).
do repeat edadp = #edadp1 #edadp2 #edadp3 #edadp4 #edadp5 #edadp6 #edadp7 #edadp8 #edadp9
  / sitacp = #sitacp1 #sitacp2 #sitacp3 #sitacp4 #sitacp5 #sitacp6 #sitacp7 #sitacp8 #sitacp9
  / adulnodep = #adulnodep1 #adulnodep2 #adulnodep3 #adulnodep4 #adulnodep5 #adulnodep6 #adulnodep7 #adulnodep8 #adulnodep9.
  compute adulnodep=$sysmis.
  do if (not(missing(edadp))).
    do if (edadp<18).
      compute adulnodep = 0 .
    end if.
  end if.
end repeat.

```

```

else if ( (edadp<25) and (sitacp=3 or sitacp=4 or sitacp=5 or sitacp=6) ) .
  compute adulnodep = 0 .
else .
  compute adulnodep = 1 .
  compute #n_adulnodep=#n_adulnodep+1.
end if.
end if.
end repeat.

vector #estcivlp(9).
do repeat estcivlp = #estcivlp1 #estcivlp2 #estcivlp3 #estcivlp4 #estcivlp5 #estcivlp6 #estcivlp7 #estcivlp8 #estcivlp9
  / v = p1_4_1 p1_4_2 p1_4_3 p1_4_4 p1_4_5 p1_4_6 p1_4_7 p1_4_8 p1_4_9 .
  compute estcivlp=$sysmis.
  if (v=1) estcivlp=1 /* 1 Soltero */.
  if (v=2 or v=3) estcivlp=2 /* 2 Casado */.
  if (v=4) estcivlp=3 /* 3 Separado */.
  if (v=5) estcivlp=5 /* 5 Divorciado */.
  if (v=6) estcivlp=4 /* 4 Viudo */.
end repeat.

compute #n_casados=0.
do repeat estcivlp = #estcivlp1 #estcivlp2 #estcivlp3 #estcivlp4 #estcivlp5 #estcivlp6 #estcivlp7 #estcivlp8 #estcivlp9.
  if (estcivlp=2) #n_casados = #n_casados+1.
end repeat.

vector #ordenp(9).
do repeat ordenp = #ordenp1 #ordenp2 #ordenp3 #ordenp4 #ordenp5 #ordenp6 #ordenp7 #ordenp8 #ordenp9
  / estcivlp = #estcivlp1 #estcivlp2 #estcivlp3 #estcivlp4 #estcivlp5 #estcivlp6 #estcivlp7 #estcivlp8 #estcivlp9.
  compute ordenp = 1.
  if (#n_casados>=2) ordenp = $sysmis .
  if (#n_casados>=2 and estcivlp >=2) ordenp = 1 .
  if (#n_casados>=2 and estcivlp =1) ordenp = 2 .
end repeat.

* ===== Obtención del SP =====.

compute cod_SP= $sysmis.
compute #aportac_SP=-1e20.

* ----- Obtención del SP si hay un único adulto no dependiente -----.

do if (#n_adulnodep=1).
  loop #i=1 to 9.
    if (#adulnodep(#i) eq 1) cod_SP= #i.
  end loop.

* ----- Obtención del SP si hay más de un adulto no dependiente -----.

else if (not #n_adulnodep=1).
  loop #i=1 to 9.
    do if (#adulnodep(#i) eq 1 and #aportac(#i) gt 0 and #ordenp(#i) eq 1 and #aportac(#i)>#aportac_SP).
      compute cod_SP= #i.
      compute #aportac_SP= #aportac(#i).
      compute #edad_SP= #edadp(#i).
    end if.
  end loop.
  loop #i=1 to 9.
    do if ( #adulnodep(#i) eq 1 and #aportac(#i) gt 0 and #ordenp(#i) eq 1 and #i ne cod_SP and #aportac(#i) =
#aportac_SP and #edadp(#i) > #edad_SP ).
      compute cod_SP= #i.
      compute #edad_SP= #edadp(#i).
      compute #sexo_SP= #sexop(#i).
    end if.
  end loop.
  loop #i=9 to 1 by -1.
    do if ( #adulnodep(#i) eq 1 and #aportac(#i) gt 0 and #ordenp(#i) eq 1 and #i ne cod_SP and #aportac(#i) =
#aportac_SP and #edadp(#i) = #edad_SP and #sexop(#i) < #sexo_SP ).
      compute cod_SP= #i.
    end if.
  end loop.

do if (missing(cod_SP)).
  loop #i=1 to 9.
    do if (#adulnodep(#i) eq 1 and #aportac(#i) gt 0 and #ordenp(#i) eq 2 and #aportac(#i)>#aportac_SP).
      compute cod_SP= #i.
      compute #aportac_SP= #aportac(#i).
      compute #edad_SP= #edadp(#i).
    end if.
  end loop.
  loop #i=1 to 9.

```

```

do if ( #adulnodep(#i) eq 1 and #aportac(#i) gt 0 and #ordenp(#i) eq 2 and #i ne cod_SP and #aportac(#i) =
#aportac_SP and #edadp(#i) > #edad_SP ).
  compute cod_SP= #i.
  compute #edad_SP= #edadp(#i).
  compute #sexo_SP= #sexop(#i).
end if.
end loop.
loop #i=9 to 1 by -1.
  do if ( #adulnodep(#i) eq 1 and #aportac(#i) gt 0 and #ordenp(#i) eq 2 and #i ne cod_SP and #aportac(#i) =
#aportac_SP and #edadp(#i) = #edad_SP and #sexop(#i) < #sexo_SP ).
    compute cod_SP= #i.
    end if.
  end loop.
end if.
end if.

do if (missing(cod_SP)).
  loop #i=1 to 9.
    do if ( #adulnodep(#i) eq 1 and ( #sitacp(#i)=1 or #sitacp(#i)=2) and #aportac(#i)>#aportac_SP).
      compute cod_SP= #i.
      compute #aportac_SP= #aportac(#i).
      compute #edad_SP= #edadp(#i).
    end if.
  end loop.
  loop #i=1 to 9.
    do if ( #adulnodep(#i) eq 1 and ( #sitacp(#i)=1 or #sitacp(#i)=2) and #i ne cod_SP and #aportac(#i) = #aportac_SP
and #edadp(#i) > #edad_SP ).
      compute cod_SP= #i.
      compute #edad_SP= #edadp(#i).
      compute #sexo_SP= #sexop(#i).
    end if.
  end loop.
  loop #i=9 to 1 by -1.
    do if ( #adulnodep(#i) eq 1 and ( #sitacp(#i)=1 or #sitacp(#i)=2) and #i ne cod_SP and #aportac(#i) = #aportac_SP
and #edadp(#i) = #edad_SP and #sexop(#i) < #sexo_SP ).
      compute cod_SP= #i.
    end if.
  end loop.
end if.

do if (missing(cod_SP)).
  loop #i=1 to 9.
    do if ( #i le nmh and #aportac(#i)>#aportac_SP).
      compute cod_SP= #i.
      compute #aportac_SP= #aportac(#i).
      compute #edad_SP= #edadp(#i).
    end if.
  end loop.
  loop #i=1 to 9.
    do if ( #i le nmh and #i ne cod_SP and #aportac(#i) = #aportac_SP and #edadp(#i) > #edad_SP ).
      compute cod_SP= #i.
      compute #edad_SP= #edadp(#i).
      compute #sexo_SP= #sexop(#i).
    end if.
  end loop.
  loop #i=9 to 1 by -1.
    do if ( #i le nmh and #i ne cod_SP and #aportac(#i) = #aportac_SP and #edadp(#i) = #edad_SP and #sexop(#i) <
#sexo_SP ).
      compute cod_SP= #i.
    end if.
  end loop.
end if.

end if.

* ===== Obtención del 2A =====.

compute cod_2A= $sysmis.
compute #aportac_2A=-1e20.

loop #i=1 to 9.
  do if ( #i ne cod_SP and #adulnodep(#i) eq 1 and #aportac(#i) gt 0 and #ordenp(#i) eq 1 and
#aportac(#i)>#aportac_2A).
    compute cod_2A= #i.
    compute #aportac_2A= #aportac(#i).
    compute #edad_2A= #edadp(#i).
  end if.
end loop.
loop #i=1 to 9.
  do if ( #i ne cod_SP and #adulnodep(#i) eq 1 and #aportac(#i) gt 0 and #ordenp(#i) eq 1 and #i ne cod_2A and
#aportac(#i) = #aportac_2A and #edadp(#i) > #edad_2A ).
    compute cod_2A= #i.
  end if.
end loop.

```

```

        compute #edad_2A= #edadp(#l).
        compute #sexo_2A= #sexop(#l).
    end if.
end loop.
loop #i=9 to 1 by -1.
    do if (#l ne cod_SP and #adulnodep(#l) eq 1 and #aportac(#l) gt 0 and #ordenp(#l) eq 1 and #l ne cod_2A and
#aportac(#l) = #aportac_2A and #edadp(#l) = #edad_2A and #sexop(#l) < #sexo_2A ).
        compute cod_2A= #l.
    end if.
end loop.

do if (missing(cod_2A)).
    loop #i=1 to 9.
        do if (#l ne cod_SP and #adulnodep(#l) eq 1 and #aportac(#l) gt 0 and #ordenp(#l) eq 2 and
#aportac(#l)>#aportac_2A).
            compute cod_2A= #l.
            compute #aportac_2A= #aportac(#l).
            compute #edad_2A= #edadp(#l).
        end if.
    end loop.
    loop #i=1 to 9.
        do if (#l ne cod_SP and #adulnodep(#l) eq 1 and #aportac(#l) gt 0 and #ordenp(#l) eq 2 and #l ne cod_2A and
#aportac(#l) = #aportac_2A and #edadp(#l) > #edad_2A ).
            compute cod_2A= #l.
            compute #edad_2A= #edadp(#l).
            compute #sexo_2A= #sexop(#l).
        end if.
    end loop.
    loop #i=9 to 1 by -1.
        do if (#l ne cod_SP and #adulnodep(#l) eq 1 and #aportac(#l) gt 0 and #ordenp(#l) eq 2 and #l ne cod_2A and
#aportac(#l) = #aportac_2A and #edadp(#l) = #edad_2A and #sexop(#l) < #sexo_2A ).
            compute cod_2A= #l.
        end if.
    end loop.
end if.

do if (missing(cod_2A)).
    loop #i=1 to 9.
        do if (#l ne cod_SP and #adulnodep(#l) eq 1 and (#sitacp(#i)=1 or #sitacp(#i)=2) and #aportac(#l)>#aportac_2A).
            compute cod_2A= #l.
            compute #aportac_2A= #aportac(#l).
            compute #edad_2A= #edadp(#l).
        end if.
    end loop.
    loop #i=1 to 9.
        do if (#l ne cod_SP and #adulnodep(#l) eq 1 and (#sitacp(#i)=1 or #sitacp(#i)=2) and #l ne cod_2A and #aportac(#l)
= #aportac_2A and #edadp(#l) > #edad_2A ).
            compute cod_2A= #l.
            compute #edad_2A= #edadp(#l).
            compute #sexo_2A= #sexop(#l).
        end if.
    end loop.
    loop #i=9 to 1 by -1.
        do if (#l ne cod_SP and #adulnodep(#l) eq 1 and (#sitacp(#i)=1 or #sitacp(#i)=2) and #l ne cod_2A and #aportac(#l)
= #aportac_2A and #edadp(#l) = #edad_2A and #sexop(#l) < #sexo_2A ).
            compute cod_2A= #l.
        end if.
    end loop.
end if.

do if (missing(cod_2A)).
    loop #i=1 to 9.
        do if ( #l le nmh and #l ne cod_SP and #aportac(#l)>#aportac_2A).
            compute cod_2A= #l.
            compute #aportac_2A= #aportac(#l).
            compute #edad_2A= #edadp(#l).
        end if.
    end loop.
    loop #i=1 to 9.
        do if ( #l le nmh and #l ne cod_SP and #l ne cod_2A and #aportac(#l) = #aportac_2A and #edadp(#l) > #edad_2A ).
            compute cod_2A= #l.
            compute #edad_2A= #edadp(#l).
            compute #sexo_2A= #sexop(#l).
        end if.
    end loop.
    loop #i=9 to 1 by -1.
        do if ( #l le nmh and #l ne cod_SP and #l ne cod_2A and #aportac(#l) = #aportac_2A and #edadp(#l) = #edad_2A
and #sexop(#l) < #sexo_2A ).
            compute cod_2A= #l.
        end if.
    end loop.
end if.

```

```

end if.

execute.

Apply Dictionary
/FROM BDH /SOURCE VARIABLES=cod_SP
/TARGET VARIABLES=cod_SP
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=cod_2A
/TARGET VARIABLES=cod_2A
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

* ----- n_percep: número de perceptores del hogar -----

vector ingr = ingr1 to ingr9 .
compute n_percep=0.
loop #i=1 to 9.
  if (ingr(#i) ne 0) n_percep = n_percep+1.
end loop.
execute.
delete variables ingr1 to ingr9.

Apply Dictionary
/FROM BDH /SOURCE VARIABLES=n_percep
/TARGET VARIABLES=n_percep
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

=====
** Carga de datos del Sustentador Principal (SP) .
=====

*----- sexo_SP: sexo del Sustentador Principal -----
compute sexo_SP=$sysmis.
do repeat k=1 2 3 4 5 6 7 8 9
  / v = p1_1_1 p1_1_2 p1_1_3 p1_1_4 p1_1_5 p1_1_6 p1_1_7 p1_1_8 p1_1_9 .
  do if (k=cod_SP) .
    if (v=1) sexo_SP =1 /* 1 Hombre */.
    if (v=2) sexo_SP =6 /* 6 Mujer */.
  end if.
end repeat.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=sexo_SP
/TARGET VARIABLES=sexo_SP
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*----- edad_SP: Edad del Sustentador Principal -----
compute edad_SP=$sysmis.
do repeat k=1 2 3 4 5 6 7 8 9
  / v = p1_2d_1 p1_2d_2 p1_2d_3 p1_2d_4 p1_2d_5 p1_2d_6 p1_2d_7 p1_2d_8 p1_2d_9 .
  do if (k=cod_SP).
    compute edad_SP =v /* edad en años */.
  end if.
end repeat.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=edad_SP
/TARGET VARIABLES=edad_SP
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*----- gedad_SP: Grupo de Edad del Sustentador Principal -----
compute gedad_SP=$sysmis.
recode edad_SP (0 thru 24=1) (25 thru 34=2) (35 thru 44=3) (45 thru 54=4) (55 thru 64=5) (65 thru 199=6) into gedad_SP.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=gedad_SP
/TARGET VARIABLES=gedad_SP
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*----- ned_SP: Nivel de Estudios del Sustentador Principal del Hogar -----
compute ned_SP=$sysmis.
do repeat k=1 2 3 4 5 6 7 8 9
  / v = p1_5_1 p1_5_2 p1_5_3 p1_5_4 p1_5_5 p1_5_6 p1_5_7 p1_5_8 p1_5_9 .
  do if (k=cod_SP) .
    if (v=1) ned_SP =1 /* 1 Sin estudios */.
    if (v=2) ned_SP =2 /* 2 Primaria */.
    if (v=3 or v=4 or v=5 ) ned_SP =3 /* 3 Secundaria, 1ª etapa */.
    if (v=6 or v=7 ) ned_SP =4 /* 4 Secundaria, 2ª etapa */.
    if (v=8 or v=9 or v=10 or v=11 or v=12 ) ned_SP =5 /* 5 Estudios Superiores y asimilados */.
  end if.
end repeat.
execute.

```

```

end if.
end repeat.
execute.
Apply Dictionary
/ FROM BDH /SOURCE VARIABLES=ned_SP
/TARGET VARIABLES=ned_SP
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*----- naci_SP: Nivel de Estudios del Sustentador Principal del Hogar (no disponible) -----,

*----- estcivil_SP: Estado Civil del Sustentador Principal del Hogar -----,
compute estcivil_SP=$sysmis.
do repeat k=1 2 3 4 5 6 7 8 9
  / v = p1_4_1 p1_4_2 p1_4_3 p1_4_4 p1_4_5 p1_4_6 p1_4_7 p1_4_8 p1_4_9 .
do if (k=cod_SP) .
  if (v=1) estcivil_SP =1          /* 1 Soltero */.
  if (v=2 or v=3) estcivil_SP =2   /* 2 Casado */.
  if (v=4) estcivil_SP =3          /* 3 Separado */.
  if (v=5) estcivil_SP =5          /* 5 Divorciado */.
  if (v=6) estcivil_SP =4          /* 4 Viudo */.
end if.
end repeat.
execute.
Apply Dictionary
/ FROM BDH /SOURCE VARIABLES=estcivil_SP
/TARGET VARIABLES=estcivil_SP
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*----- situac2_SP: Situacion Actividad del Sustentador Principal del Hogar -----,
compute situac2_SP=$sysmis.
if (gedad_SP=0) situac2_SP=0          /* 0 ND (menor de 16 años) */.
do repeat h_tca = h_tca_1 h_tca_2 h_tca_3 h_tca_4 h_tca_5 h_tca_6 h_tca_7 h_tca_8 h_tca_9 .
  compute h_tca=0.
end repeat.
do repeat v1 = p6_12_1_1 p6_12_2_1 p6_12_3_1 p6_12_4_1 p6_12_5_1 p6_12_6_1 p6_12_7_1 p6_12_8_1 p6_12_9_1
  /v2 = p6_12_1_2 p6_12_2_2 p6_12_3_2 p6_12_4_2 p6_12_5_2 p6_12_6_2 p6_12_7_2 p6_12_8_2 p6_12_9_2
  /v3 = p6_12_1_3 p6_12_2_3 p6_12_3_3 p6_12_4_3 p6_12_5_3 p6_12_6_3 p6_12_7_3 p6_12_8_3 p6_12_9_3
  /h_tca = h_tca_1 h_tca_2 h_tca_3 h_tca_4 h_tca_5 h_tca_6 h_tca_7 h_tca_8 h_tca_9 .
  if (v1>0) h_tca=h_tca+v1.
  if (v2>0) h_tca=h_tca+v2.
  if (v3>0) h_tca=h_tca+v3.
end repeat.
execute.
do repeat h_tcp = h_tcp_1 h_tcp_2 h_tcp_3 h_tcp_4 h_tcp_5 h_tcp_6 h_tcp_7 h_tcp_8 h_tcp_9 .
  compute h_tcp=0.
end repeat.
do repeat v1 = p6_33_1_1 p6_33_2_1 p6_33_3_1 p6_33_4_1 p6_33_5_1 p6_33_6_1 p6_33_7_1 p6_33_8_1 p6_33_9_1
  /v2 = p6_33_1_2 p6_33_2_2 p6_33_3_2 p6_33_4_2 p6_33_5_2 p6_33_6_2 p6_33_7_2 p6_33_8_2 p6_33_9_2
  /v3 = p6_33_1_3 p6_33_2_3 p6_33_3_3 p6_33_4_3 p6_33_5_3 p6_33_6_3 p6_33_7_3 p6_33_8_3 p6_33_9_3
  /h_tcp = h_tcp_1 h_tcp_2 h_tcp_3 h_tcp_4 h_tcp_5 h_tcp_6 h_tcp_7 h_tcp_8 h_tcp_9 .
  if (v1>0) h_tcp=h_tcp+v1.
  if (v2>0) h_tcp=h_tcp+v2.
  if (v3>0) h_tcp=h_tcp+v3.
end repeat.
execute.
do repeat k=1 2 3 4 5 6 7 8 9
  / v1 = p6_1c1_1 p6_1c1_2 p6_1c1_3 p6_1c1_4 p6_1c1_5 p6_1c1_6 p6_1c1_7 p6_1c1_8 p6_1c1_9
  / v2 = p6_1c2_1 p6_1c2_2 p6_1c2_3 p6_1c2_4 p6_1c2_5 p6_1c2_6 p6_1c2_7 p6_1c2_8 p6_1c2_9
  / v3 = p6_1c3_1 p6_1c3_2 p6_1c3_3 p6_1c3_4 p6_1c3_5 p6_1c3_6 p6_1c3_7 p6_1c3_8 p6_1c3_9
  / v4 = p6_1c4_1 p6_1c4_2 p6_1c4_3 p6_1c4_4 p6_1c4_5 p6_1c4_6 p6_1c4_7 p6_1c4_8 p6_1c4_9
  / v5 = p6_1c5_1 p6_1c5_2 p6_1c5_3 p6_1c5_4 p6_1c5_5 p6_1c5_6 p6_1c5_7 p6_1c5_8 p6_1c5_9
  / v6 = p6_1c6_1 p6_1c6_2 p6_1c6_3 p6_1c6_4 p6_1c6_5 p6_1c6_6 p6_1c6_7 p6_1c6_8 p6_1c6_9
  / v7 = p6_1c7_1 p6_1c7_2 p6_1c7_3 p6_1c7_4 p6_1c7_5 p6_1c7_6 p6_1c7_7 p6_1c7_8 p6_1c7_9
  / v8 = p6_1c8_1 p6_1c8_2 p6_1c8_3 p6_1c8_4 p6_1c8_5 p6_1c8_6 p6_1c8_7 p6_1c8_8 p6_1c8_9
  / h_tca = h_tca_1 h_tca_2 h_tca_3 h_tca_4 h_tca_5 h_tca_6 h_tca_7 h_tca_8 h_tca_9
  / h_tcp = h_tcp_1 h_tcp_2 h_tcp_3 h_tcp_4 h_tcp_5 h_tcp_6 h_tcp_7 h_tcp_8 h_tcp_9 .
do if (k=cod_SP) .
  do if ((v1=1 or v2=1) and h_tcp<h_tca and h_tca>=30).
    compute situac2_SP =1          /* 1 Asalariado a tiempo completo */.
  else if ((v1=1 or v2=1) and h_tcp<h_tca and h_tca<30).
    compute situac2_SP =2          /* 2 Asalariado a tiempo parcial */.
  else if ((v1=1 or v2=1) and h_tcp>=h_tca and h_tcp>=30).
    compute situac2_SP =3          /* 3 Trabajador por cuenta propia a tiempo completo */.
  else if ((v1=1 or v2=1) and h_tcp>=h_tca and h_tcp<30).
    compute situac2_SP =4          /* 4 Trabajador por cuenta propia a tiempo parcial */.
  else if (v4=1).
    compute situac2_SP =7          /* 7 Jubilado, retirado, jubilado anticipado o ha cerrado un negocio */.
  else if (v5=1).
    compute situac2_SP =8          /* 8 Incapacitado permanente para trabajar */.
  else if (v6=1).
    compute situac2_SP =6          /* 6 Estudiante, escolar o en formación */.

```



```

else if (v3=1).
  compute situac2_SP =5 /* 5 Parado */.
else if (v7=1).
  compute situac2_SP =10 /* 10 Dedicado a las labores del hogar, al cuidado de niños u otras personas */.
else if (v8=1).
  compute situac2_SP =11 /* 11 Otro clase de inactividad económica */.
end if.
end if.
end repeat.
execute.
* ---- 9 Servicio militar obligatorio o prestación social sustitutoria (no existe en la EFF) .
delete variables h_tca_1 to h_tcp_9.
Apply Dictionary
/ FROM BDH /SOURCE VARIABLES=situac2_SP
/TARGET VARIABLES=situac2_SP
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*----- ocupac_SP: Ocupación del Sustentador Principal del Hogar -----.
compute ocupac_SP=$sysmis.
do repeat k=1 2 3 4 5 6 7 8 9
  / v = p6_3_1 p6_3_2 p6_3_3 p6_3_4 p6_3_5 p6_3_6 p6_3_7 p6_3_8 p6_3_9 .
do if (k=cod_SP) .
  if (v=1) ocupac_SP =1 /* 1 Dirección de las empresas y de la administración pública */.
  if (v=2) ocupac_SP =2 /* 2 Técnicos y profesionales científicos e intelectuales */.
  if (v=3) ocupac_SP =3 /* 3 Técnicos y profesionales de apoyo */.
  if (v=4) ocupac_SP =4 /* 4 Empleados de tipo administrativo */.
  if (v=5) ocupac_SP =5 /* 5 Trabajadores de servicios de restauración, personales... */.
  if (v=6) ocupac_SP =6 /* 6 Trabajadores cualificados en la agricultura y en la pesca */.
  if (v=7) ocupac_SP =7 /* 7 Artesanos y trabajadores cualificados de las industrias ... */.
  if (v=8) ocupac_SP =8 /* 8 Operadores y montadores de instalaciones y maquinaria ... */.
  if (v=9) ocupac_SP =9 /* 9 Trabajadores no cualificados */.
  if (v=10) ocupac_SP =10 /* 10 Fuerzas armadas */.
end if.
end repeat.
if (situac2_SP>=5 and situac2_SP<=11) ocupac_SP= -2 /* _NA No Aplicable - definido en la ECV sólo para ocupados */.
if (gedad_SP=0) ocupac_SP=0 /* 0 ND (menor de 16 años) */.
execute.
Apply Dictionary
/ FROM BDH /SOURCE VARIABLES=ocupac_SP
/TARGET VARIABLES=ocupac_SP
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

lif (leff='EFF08') !then.
*----- activestb1_SP: Actividad del Establecimiento de Trabajo del Sustentador Principal del Hogar (1ª Clasif)-----.
String activestb1_SP (A5).
compute activestb1_SP=''.
do repeat k=1 2 3 4 5 6 7 8 9
  / v = p6_4_1 p6_4_2 p6_4_3 p6_4_4 p6_4_5 p6_4_6 p6_4_7 p6_4_8 p6_4_9 .
do if (k=cod_SP) .
  if (v=1) activestb1_SP ='A+B' /* A+B Agricultura, ganadería, caza, silvicultura y pesca */.
  if (v=2 or v=3 or v=4) activestb1_SP ='C+D+E' /* C+D+E Indus. extractivas y manufact. Prod. y Distrib. de energía eléctrica, gas y agua */.
  if (v=5) activestb1_SP ='F' /* F Construcción */.
  if (v=6) activestb1_SP ='G' /* G Comercio; reparación de vehículos de motor, ciclomotores y artículos de uso doméstico */.
  if (v=7) activestb1_SP ='H' /* H Hostelería */.
  if (v=8) activestb1_SP ='I' /* I Transporte, almacenamiento y comunicaciones */.
  if (v=9) activestb1_SP ='J' /* J Intermediación financiera */.
  if (v=10) activestb1_SP ='K' /* K Actividades inmobiliarias y de alquiler; Servicios empresariales */.
  if (v=11) activestb1_SP ='L' /* L Administración pública y defensa; seguridad social obligatoria */.
  if (v=12) activestb1_SP ='M' /* M Educación */.
  if (v=13) activestb1_SP ='N' /* N Actividades sanitarias y veterinarias; servicios sociales */.
  if (v=14 or v=15 or v=16) activestb1_SP ='O+P+Q' /* O+P+Q Otras Activ y de servicios a la comunidad ; Serv personales. Hogares que emplea p.doméstico. Organ extraterritoriales */.
end if.
end repeat.
if (situac2_SP>=5 and situac2_SP<=11) activestb1_SP='_NA' /* _NA No Aplicable - definido en la ECV sólo para ocupados */.
if (gedad_SP=0) activestb1_SP='_ND' /* _ND ND (menor de 16 años) */.
execute.
Apply Dictionary
/ FROM BDH /SOURCE VARIABLES=activestb1_SP
/TARGET VARIABLES=activestb1_SP
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
lifend.

lif (leff='EFF11') !then.
*----- activestb2_SP: Actividad del Establecimiento de Trabajo del Sustentador Principal del Hogar -----.
String activestb2_SP (A3).
compute activestb2_SP=''.
do repeat k=1 2 3 4 5 6 7 8 9

```

```

/ v = p6_4_1 p6_4_2 p6_4_3 p6_4_4 p6_4_5 p6_4_6 p6_4_7 p6_4_8 p6_4_9 .
do if (k=cod_SP) .
  if (v=1) activestb2_SP ='A'          /* A Agricultura, ganadería, silvicultura y pesca */.
  if (v=2) activestb2_SP ='B'          /* B Industrias extractivas */.
  if (v=3) activestb2_SP ='C'          /* C Industria manufacturera */.
  if (v=4) activestb2_SP ='D'          /* D Suministro de energía eléctrica, gas, vapor y aire ac ... */.
  if (v=5) activestb2_SP ='E'          /* E Suministro de agua, actividades de saneamiento, ... */.
  if (v=6) activestb2_SP ='F'          /* F Construcción */.
  if (v=7) activestb2_SP ='G'          /* G Comercio al por mayor y al por menor; reparación ... */.
  if (v=8) activestb2_SP ='H'          /* H Transporte y almacenamiento */.
  if (v=9) activestb2_SP ='I'          /* I Hostelería */.
  if (v=10) activestb2_SP ='J'         /* J Información y comunicaciones */.
  if (v=11) activestb2_SP ='K'         /* K Actividades financieras y de seguros */.
  if (v=12) activestb2_SP ='L'         /* L Actividades inmobiliarias */.
  if (v=13) activestb2_SP ='M'         /* M Actividades profesionales, científicas y técnicas */.
  if (v=14) activestb2_SP ='N'         /* N Actividades administrativas y servicios auxiliares */.
  if (v=15) activestb2_SP ='O'         /* O Administración pública y defensa; seguridad social obligatoria */.
  if (v=16) activestb2_SP ='P'         /* P Educación */.
  if (v=17) activestb2_SP ='Q'         /* Q Actividades sanitarias y de servicios sociales */.
  if (v=18) activestb2_SP ='R'         /* R Actividades artísticas, recreativas y de entretenimiento */.
  if (v=19) activestb2_SP ='S'         /* S Otros servicios */.
  if (v=20) activestb2_SP ='T'         /* T Actividades de los hogares como empleadores doméstico */.
  if (v=21) activestb2_SP ='U'         /* U Actividades de organizaciones y organismos extraterritoriales */.
end if.
end repeat.
if (situac2_SP>=5 and situac2_SP<=11) activestb2_SP='_NA'          /* _NA No Aplicable - definido en la ECV sólo para ocupados */.
if (gedad_SP=0) activestb2_SP='_ND'          /* _ND ND (menor de 16 años) */.
execute.
Apply Dictionary
/ FROM BDH /SOURCE VARIABLES=activestb2_SP
/TARGET VARIABLES=activestb2_SP
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
lifend.

lif (leff='EFF08') !then.
*---- activestb1_SP_m: Actividad del establecimiento laboral (1ª clasif) -----
RECODE activestb1_SP (" '=SYSMIS) ('_NA'=-2) ('_ND'=0) ('A+B'=1) ('C+D+E'=2) ('F'=3) ('G'=4) ('H'=5) ('I'=6) ('J'=7)
('K'=8) ('L'=9) ('M'=10) ('N'=11) ('O+P+Q'=12) INTO activestb1_SP_m.
execute.
Apply Dictionary
/ FROM BDH /SOURCE VARIABLES=activestb1_SP_m
/TARGET VARIABLES=activestb1_SP_m
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
lifend.

lif (leff='EFF11') !then.
*---- activestb2_SP_m: Actividad del establecimiento laboral (2ª clasif) -----
RECODE activestb2_SP (" '=SYSMIS) ('_NA'=-2) ('_ND'=0) ('A'=1) ('B'=2) ('C'=3) ('D'=4) ('E'=5) ('F'=6) ('G'=7)
('H'=8) ('I'=9) ('J'=10) ('K'=11) ('L'=12) ('M'=13) ('N'=14) ('O'=15) ('P'=16)
('Q'=17) ('R'=18) ('S'=19) ('T'=20) ('U'=21) INTO activestb2_SP_m.
execute.
Apply Dictionary
/ FROM BDH /SOURCE VARIABLES=activestb2_SP_m
/TARGET VARIABLES=activestb2_SP_m
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
lifend.

*----- situprof_SP: Situación Profesional del Sustentador Principal del Hogar -----
compute situprof_SP=$sysmis.
lif (leff='EFF08') !then.
do repeat k=1 2 3 4 5 6
  / v1 = p4_103s1_1 p4_103s1_2 p4_103s1_3 p4_103s1_4 p4_103s1_5 p4_103s1_6
  / v2 = p4_103s2_1 p4_103s2_2 p4_103s2_3 p4_103s2_4 p4_103s2_5 p4_103s2_6
  / v3 = p4_103s3_1 p4_103s3_2 p4_103s3_3 p4_103s3_4 p4_103s3_5 p4_103s3_6
  / v4 = p4_103s4_1 p4_103s4_2 p4_103s4_3 p4_103s4_4 p4_103s4_5 p4_103s4_6
  / w1 = p4_109_1 p4_109_2 p4_109_3 p4_109_4 p4_109_5 p4_109_6 .
do if (situac2_SP=3 or situac2_SP=4) .
  do if (k=1).
    if ((v1=cod_SP or v2=cod_SP or v3=cod_SP or v4=cod_SP) and (w1 - 1 >1)) situprof_SP=1          /* 1 Empleador */.
    if ((v1=cod_SP or v2=cod_SP or v3=cod_SP or v4=cod_SP) and (w1 - 1 =1)) situprof_SP=2          /* 2 Empresario
sin asalariados o trabajador independiente */.
  else if (k>1).
    if ((v1=cod_SP or v2=cod_SP or v3=cod_SP or v4=cod_SP) and (w1 - 1 >1)) situprof_SP=min(situprof_SP,1)          /*
1 Empleador */.
    if ((v1=cod_SP or v2=cod_SP or v3=cod_SP or v4=cod_SP) and (w1 - 1 =1)) situprof_SP=min(situprof_SP,2)          /*
2 Empresario sin asalariados o trabajador independiente */.
  end if.
end if.
end repeat.
lifend

```

```

!if (!eff='EFF11') !then.
do repeat k=1 2 3 4 5 6
  / v1 = p4_103s1_1 p4_103s1_2 p4_103s1_3 p4_103s1_4 p4_103s1_5 p4_103s1_6
  / v2 = p4_103s2_1 p4_103s2_2 p4_103s2_3 p4_103s2_4 p4_103s2_5 p4_103s2_6
  / v3 = p4_103s3_1 p4_103s3_2 p4_103s3_3 p4_103s3_4 p4_103s3_5 p4_103s3_6
  / v4 = p4_103s4_1 p4_103s4_2 p4_103s4_3 p4_103s4_4 p4_103s4_5 p4_103s4_6
  / w1 = p4_109_1 p4_109_2 p4_109_3 p4_109_4 p4_109_5 p4_109_6
  / w2 = p4_109a_1 p4_109a_2 p4_109a_3 p4_109a_4 p4_109a_5 p4_109a_6 .
do if (situac2_SP=3 or situac2_SP=4) .
  do if (k=1).
    if ((v1=cod_SP or v2=cod_SP or v3=cod_SP or v4=cod_SP) and (w1 - 1*(w2=1)>1) ) situprof_SP=1 /* 1
Empleador */.
    if ((v1=cod_SP or v2=cod_SP or v3=cod_SP or v4=cod_SP) and (w1 - 1*(w2=1)=1) ) situprof_SP =2 /* 2
Empresario sin asalariados o trabajador independiente */.
    else if (k>1).
      if ((v1=cod_SP or v2=cod_SP or v3=cod_SP or v4=cod_SP) and (w1 - 1*(w2=1)>1) ) situprof_SP=min(situprof_SP,1)
/* 1 Empleador */.
      if ((v1=cod_SP or v2=cod_SP or v3=cod_SP or v4=cod_SP) and (w1 - 1*(w2=1)=1) ) situprof_SP=min(situprof_SP,2)
/* 2 Empresario sin asalariados o trabajador independiente */.
    end if.
  end if.
end repeat.
!lifend

if (situac2_SP=1 or situac2_SP=2) situprof_SP=3 /* 3 Asalariado */.
/* 4 Otra situación (Ayuda Familiar) */.
if (situac2_SP>=5 and situac2_SP<=11) situprof_SP=-2 /* -2 No Aplicable - definido en la ECV sólo para ocupados */.
if (gedad_SP=0) situprof_SP=0 /* 0 ND (menor de 16 años) */.
execute.

Apply Dictionary
/ FROM BDH /SOURCE VARIABLES=situprof_SP
/TARGET VARIABLES=situprof_SP
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*----- tipcont_SP: Tipo de contrato del Sustentador Principal del Hogar -----.
compute tipcont_SP=$sysmis.
do repeat k=1 2 3 4 5 6 7 8 9
  / v1 = p6_13_1_1 p6_13_2_1 p6_13_3_1 p6_13_4_1 p6_13_5_1 p6_13_6_1 p6_13_7_1 p6_13_8_1 p6_13_9_1
  / v2 = p6_13_1_2 p6_13_2_2 p6_13_3_2 p6_13_4_2 p6_13_5_2 p6_13_6_2 p6_13_7_2 p6_13_8_2 p6_13_9_2
  / v3 = p6_13_1_3 p6_13_2_3 p6_13_3_3 p6_13_4_3 p6_13_5_3 p6_13_6_3 p6_13_7_3 p6_13_8_3 p6_13_9_3 .
do if (k=cod_SP) .
  if (v1=2 or v2=2 or v3=2) tipcont_SP =2 /* 2 Contrato temporal de duración determinada */.
  if (v1=1 or v2=1 or v3=1) tipcont_SP =1 /* 1 Contrato fijo de duración indefinida */.
end if.
end repeat.
if (situac2_SP>=3 and situac2_SP<=11) tipcont_SP=-2 /* -2 No asalariado (No Aplicable) */.
if (gedad_SP=0) tipcont_SP=0 /* 0 ND (menor de 16 años) */.
execute.

Apply Dictionary
/ FROM BDH /SOURCE VARIABLES=tipcont_SP
/TARGET VARIABLES=tipcont_SP
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

!if (!eff='EFF08') !then.
*----- cse_SP: Categoría socio-económica del Sustentador Principal del Hogar -----.
compute cse_SP=$sysmis.
do if (situprof_SP =1).
  do if (activestb1_SP ='A+B' or ocupac_SP=6) .
    compute cse_SP=1 /* 1 Empres. Agrar. c/Asal */.
  else if (activestb1_SP ge 'C' and activestb1_SP le 'Q') .
    compute cse_SP=2 /* 2 Empres. No Agr. c/Asal */.
  else .
    compute cse_SP=3 /* 3 Otros Empresarios c/Asal No clasif */.
  end if.
else if (situprof_SP =2).
  do if (activestb1_SP ='A+B' or ocupac_SP=6) .
    compute cse_SP=4 /* 4 Empres. Agrar. s/Asal */.
  else if (activestb1_SP ge 'C' and activestb1_SP le 'Q') .
    compute cse_SP=5 /* 5 Empres. No Agr. s/Asal */.
  else .
    compute cse_SP=6 /* 6 Otros Empresarios s/Asal No clasif */.
  end if.
else if (situprof_SP =3).
  do if (activestb1_SP ='A+B' or ocupac_SP=6) .
    compute cse_SP=7 /* 7 Asalariados Agrarios */.
  else if (activestb1_SP ge 'C' and activestb1_SP le 'Q') .
    do if (ocupac_SP=1 or ocupac_SP=2).
      compute cse_SP=8 /* 8 Cuadros Super. No Agr. */.
    end if.
  end if.
end if.

```

```

else if (ocupac_SP>=3 and ocupac_SP<=8).
  compute cse_SP=9 /* 9 Resto Empleados cualif. No Agr. */.
else if (ocupac_SP=9) .
  compute cse_SP=10 /* 10 Empleados no cualif. No Agrar. */.
else if (ocupac_SP=10) .
  compute cse_SP=11 /* 11 Fuerzas Armadas */.
end if.
else .
  compute cse_SP=12 /* 12 Otros Asalariados No clasif */.
end if.
else if (situac2_SP=5) .
  compute cse_SP=13 /* 13 Parados */.
else if (situac2_SP=7) .
  compute cse_SP=15 /* 15 Retirado */.
else if (situac2_SP=6) .
  compute cse_SP=14 /* 14 Formación */.
else if (situac2_SP=8 or situac2_SP=10 or situac2_SP=11) .
  compute cse_SP=16 /* 16 Otros Inactivos */.
end if.
if (gedad_SP=0) cse_SP=0 /* 0 ND (menor de 16 años) */.
lifend.
!if (!eff='EFF11') !then.
*----- cse_SP: Categoría socio-económica del Sustentador Principal del Hogar -----
compute cse_SP=$sysmis.
do if (situprof_SP=1) .
  do if (activestb2_SP='A' or ocupac_SP=6) .
    compute cse_SP=1 /* 1 Empres. Agrar. c/Asal */.
  else if (activestb2_SP ge 'B' and activestb2_SP le 'U') .
    compute cse_SP=2 /* 2 Empres. No Agr. c/Asal */.
  else .
    compute cse_SP=3 /* 3 Otros Empresarios c/Asal No clasif */.
  end if.
else if (situprof_SP=2) .
  do if (activestb2_SP='A' or ocupac_SP=6) .
    compute cse_SP=4 /* 4 Empres. Agrar. s/Asal */.
  else if (activestb2_SP ge 'B' and activestb2_SP le 'U') .
    compute cse_SP=5 /* 5 Empres. No Agr. s/Asal */.
  else .
    compute cse_SP=6 /* 6 Otros Empresarios s/Asal No clasif */.
  end if.
else if (situprof_SP=3) .
  do if (activestb2_SP='A' or ocupac_SP=6) .
    compute cse_SP=7 /* 7 Asalariados Agrarios */.
  else if (activestb2_SP ge 'B' and activestb2_SP le 'U') .
    do if (ocupac_SP=1 or ocupac_SP=2) .
      compute cse_SP=8 /* 8 Cuadros Super. No Agr. */.
    else if (ocupac_SP>=3 and ocupac_SP<=8) .
      compute cse_SP=9 /* 9 Resto Empleados cualif. No Agr. */.
    else if (ocupac_SP=9) .
      compute cse_SP=10 /* 10 Empleados no cualif. No Agrar. */.
    else if (ocupac_SP=10) .
      compute cse_SP=11 /* 11 Fuerzas Armadas */.
    end if.
  else .
    compute cse_SP=12 /* 12 Otros Asalariados No clasif */.
  end if.
else if (situac2_SP=5) .
  compute cse_SP=13 /* 13 Parados */.
else if (situac2_SP=7) .
  compute cse_SP=15 /* 15 Retirado */.
else if (situac2_SP=6) .
  compute cse_SP=14 /* 14 Formación */.
else if (situac2_SP=8 or situac2_SP=10 or situac2_SP=11) .
  compute cse_SP=16 /* 16 Otros Inactivos */.
end if.
if (gedad_SP=0) cse_SP=0 /* 0 ND (menor de 16 años) */.
lifend.
execute.

Apply Dictionary
/FROM BDH /SOURCE VARIABLES=cse_SP
/TARGET VARIABLES=cse_SP
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*=====
** Carga de Datos del 2º Adulto en importancia de su aportacion al Hogar (2A) .
*=====
*----- sexo_2A: sexo del 2º Adulto -----
compute sexo_2A=$sysmis.
do repeat k=1 2 3 4 5 6 7 8 9
  /v = p1_1_1 p1_1_2 p1_1_3 p1_1_4 p1_1_5 p1_1_6 p1_1_7 p1_1_8 p1_1_9 .

```

```

do if (k=cod_2A) .
  if (v=1) sexo_2A =1 /* 1 Hombre */.
  if (v=2) sexo_2A =6 /* 6 Mujer */.
end if.
end repeat.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=sexo_2A
/TARGET VARIABLES=sexo_2A
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*----- edad_2A: Edad del 2° Adulto -----,
compute edad_2A=$sysmis.
do repeat k=1 2 3 4 5 6 7 8 9
  / v = p1_2d_1 p1_2d_2 p1_2d_3 p1_2d_4 p1_2d_5 p1_2d_6 p1_2d_7 p1_2d_8 p1_2d_9 .
  do if (k=cod_2A).
    compute edad_2A =v /* edad en años */.
  end if.
end repeat.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=edad_2A
/TARGET VARIABLES=edad_2A
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*----- gedad_2A: Grupo de Edad del 2° Adulto -----,
compute gedad_2A=$sysmis.
recode edad_2A (0 thru 24=1) (25 thru 34=2) (35 thru 44=3) (45 thru 54=4) (55 thru 64=5) (65 thru 199=6) into gedad_2A.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=gedad_2A
/TARGET VARIABLES=gedad_2A
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*----- ned_2A: Nivel de Estudios del 2° Adulto del Hogar -----,
compute ned_2A=$sysmis.
do repeat k=1 2 3 4 5 6 7 8 9
  / v = p1_5_1 p1_5_2 p1_5_3 p1_5_4 p1_5_5 p1_5_6 p1_5_7 p1_5_8 p1_5_9 .
  do if (k=cod_2A) .
    if (v=1) ned_2A =1 /* 1 Sin estudios */.
    if (v=2) ned_2A =2 /* 2 Primaria */.
    if (v=3 or v=4 or v=5 ) ned_2A =3 /* 3 Secundaria, 1ª etapa */.
    if (v=6 or v=7 ) ned_2A =4 /* 4 Secundaria, 2ª etapa */.
    if (v=8 or v=9 or v=10 or v=11 or v=12 ) ned_2A =5 /* 5 Estudios Superiores y asimilados */.
  end if.
end repeat.
execute.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=ned_2A
/TARGET VARIABLES=ned_2A
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*----- naci_2A: Nivel de Estudios del 2° Adulto del Hogar (no disponible) -----,

*----- estcivil_2A: Estado Civil del 2° Adulto del Hogar -----,
compute estcivil_2A=$sysmis.
do repeat k=1 2 3 4 5 6 7 8 9
  / v = p1_4_1 p1_4_2 p1_4_3 p1_4_4 p1_4_5 p1_4_6 p1_4_7 p1_4_8 p1_4_9 .
  do if (k=cod_2A) .
    if (v=1) estcivil_2A =1 /* 1 Soltero */.
    if (v=2 or v=3) estcivil_2A =2 /* 2 Casado */.
    if (v=4) estcivil_2A =3 /* 3 Separado */.
    if (v=5) estcivil_2A =5 /* 5 Divorciado */.
    if (v=6) estcivil_2A =4 /* 4 Viudo */.
  end if.
end repeat.
Apply Dictionary
/FROM BDH /SOURCE VARIABLES=estcivil_2A
/TARGET VARIABLES=estcivil_2A
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*----- situac2_2A: Situación Actividad del 2° Adulto del Hogar -----,
compute situac2_2A=$sysmis.
if (gedad_2A=0) situac2_2A=0 /* 0 ND (menor de 16 años) */.
do repeat h_tca = h_tca_1 h_tca_2 h_tca_3 h_tca_4 h_tca_5 h_tca_6 h_tca_7 h_tca_8 h_tca_9 .
  compute h_tca=0.
end repeat.
do repeat v1 = p6_12_1_1 p6_12_2_1 p6_12_3_1 p6_12_4_1 p6_12_5_1 p6_12_6_1 p6_12_7_1 p6_12_8_1 p6_12_9_1
  /v2 = p6_12_1_2 p6_12_2_2 p6_12_3_2 p6_12_4_2 p6_12_5_2 p6_12_6_2 p6_12_7_2 p6_12_8_2 p6_12_9_2
  /v3 = p6_12_1_3 p6_12_2_3 p6_12_3_3 p6_12_4_3 p6_12_5_3 p6_12_6_3 p6_12_7_3 p6_12_8_3 p6_12_9_3
  /h_tca = h_tca_1 h_tca_2 h_tca_3 h_tca_4 h_tca_5 h_tca_6 h_tca_7 h_tca_8 h_tca_9 .

```

```

if (v1>0) h_tca=h_tca+v1.
if (v2>0) h_tca=h_tca+v2.
if (v3>0) h_tca=h_tca+v3.
end repeat.
execute.
do repeat h_tcp = h_tcp_1 h_tcp_2 h_tcp_3 h_tcp_4 h_tcp_5 h_tcp_6 h_tcp_7 h_tcp_8 h_tcp_9 .
  compute h_tcp=0.
end repeat.
do repeat v1 = p6_33_1_1 p6_33_2_1 p6_33_3_1 p6_33_4_1 p6_33_5_1 p6_33_6_1 p6_33_7_1 p6_33_8_1 p6_33_9_1
  /v2 = p6_33_1_2 p6_33_2_2 p6_33_3_2 p6_33_4_2 p6_33_5_2 p6_33_6_2 p6_33_7_2 p6_33_8_2 p6_33_9_2
  /v3 = p6_33_1_3 p6_33_2_3 p6_33_3_3 p6_33_4_3 p6_33_5_3 p6_33_6_3 p6_33_7_3 p6_33_8_3 p6_33_9_3
  /h_tcp = h_tcp_1 h_tcp_2 h_tcp_3 h_tcp_4 h_tcp_5 h_tcp_6 h_tcp_7 h_tcp_8 h_tcp_9 .
  if (v1>0) h_tcp=h_tcp+v1.
  if (v2>0) h_tcp=h_tcp+v2.
  if (v3>0) h_tcp=h_tcp+v3.
end repeat.
execute.
do repeat k=1 2 3 4 5 6 7 8 9
  /v1 = p6_1c1_1 p6_1c1_2 p6_1c1_3 p6_1c1_4 p6_1c1_5 p6_1c1_6 p6_1c1_7 p6_1c1_8 p6_1c1_9
  /v2 = p6_1c2_1 p6_1c2_2 p6_1c2_3 p6_1c2_4 p6_1c2_5 p6_1c2_6 p6_1c2_7 p6_1c2_8 p6_1c2_9
  /v3 = p6_1c3_1 p6_1c3_2 p6_1c3_3 p6_1c3_4 p6_1c3_5 p6_1c3_6 p6_1c3_7 p6_1c3_8 p6_1c3_9
  /v4 = p6_1c4_1 p6_1c4_2 p6_1c4_3 p6_1c4_4 p6_1c4_5 p6_1c4_6 p6_1c4_7 p6_1c4_8 p6_1c4_9
  /v5 = p6_1c5_1 p6_1c5_2 p6_1c5_3 p6_1c5_4 p6_1c5_5 p6_1c5_6 p6_1c5_7 p6_1c5_8 p6_1c5_9
  /v6 = p6_1c6_1 p6_1c6_2 p6_1c6_3 p6_1c6_4 p6_1c6_5 p6_1c6_6 p6_1c6_7 p6_1c6_8 p6_1c6_9
  /v7 = p6_1c7_1 p6_1c7_2 p6_1c7_3 p6_1c7_4 p6_1c7_5 p6_1c7_6 p6_1c7_7 p6_1c7_8 p6_1c7_9
  /v8 = p6_1c8_1 p6_1c8_2 p6_1c8_3 p6_1c8_4 p6_1c8_5 p6_1c8_6 p6_1c8_7 p6_1c8_8 p6_1c8_9
  /h_tca = h_tca_1 h_tca_2 h_tca_3 h_tca_4 h_tca_5 h_tca_6 h_tca_7 h_tca_8 h_tca_9
  /h_tcp = h_tcp_1 h_tcp_2 h_tcp_3 h_tcp_4 h_tcp_5 h_tcp_6 h_tcp_7 h_tcp_8 h_tcp_9 .
  do if (k=cod_2A) .
    do if ((v1=1 or v2=1) and h_tcp<h_tca and h_tca>=30).
      compute situac2_2A =1 /* 1 Asalariado a tiempo completo */ .
    else if ((v1=1 or v2=1) and h_tcp<h_tca and h_tca<30).
      compute situac2_2A =2 /* 2 Asalariado a tiempo parcial */ .
    else if ((v1=1 or v2=1) and h_tcp>=h_tca and h_tcp>=30).
      compute situac2_2A =3 /* 3 Trabajador por cuenta propia a tiempo completo */.
    else if ((v1=1 or v2=1) and h_tcp>=h_tca and h_tcp<30).
      compute situac2_2A =4 /* 4 Trabajador por cuenta propia a tiempo parcial */ .
    else if (v4=1).
      compute situac2_2A =7 /* 7 Jubilado, retirado, jubilado anticipado o ha cerrado un negocio */.
    else if (v5=1).
      compute situac2_2A =8 /* 8 Incapacitado permanente para trabajar */.
    else if (v6=1).
      compute situac2_2A =6 /* 6 Estudiante, escolar o en formación */.
    else if (v3=1).
      compute situac2_2A =5 /* 5 Parado */ .
    else if (v7=1).
      compute situac2_2A =10 /* 10 Dedicado a las labores del hogar, al cuidado de niños ... */ .
    else if (v8=1).
      compute situac2_2A =11 /* 11 Otro clase de inactividad económica */ .
    end if.
  end if.
end repeat.
* ---- 9 Servicio militar obligatorio o prestación social sustitutoria (no existe en la EFF) .
execute.
delete variables h_tca_1 to h_tcp_9.
Apply Dictionary
/ FROM BDH /SOURCE VARIABLES=situac2_2A
/TARGET VARIABLES=situac2_2A
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*----- ocupac_2A: Ocupación del 2º Adulto del Hogar -----.
compute ocupac_2A=$sysmis.
do repeat k=1 2 3 4 5 6 7 8 9
  /v = p6_3_1 p6_3_2 p6_3_3 p6_3_4 p6_3_5 p6_3_6 p6_3_7 p6_3_8 p6_3_9 .
  do if (k=cod_2A) .
    if (v=1) ocupac_2A =1 /* 1 Dirección de las empresas y de la administración pública */.
    if (v=2) ocupac_2A =2 /* 2 Técnicos y profesionales científicos e intelectuales */.
    if (v=3) ocupac_2A =3 /* 3 Técnicos y profesionales de apoyo */.
    if (v=4) ocupac_2A =4 /* 4 Empleados de tipo administrativo */.
    if (v=5) ocupac_2A =5 /* 5 Trabajadores de servicios de restauración, personales, ... */.
    if (v=6) ocupac_2A =6 /* 6 Trabajadores cualificados en la agricultura y en la pesca */.
    if (v=7) ocupac_2A =7 /* 7 Artesanos y trabajadores cualificados de las industrias ... */.
    if (v=8) ocupac_2A =8 /* 8 Operadores y montadores de instalaciones y maquinaria ... */.
    if (v=9) ocupac_2A =9 /* 9 Trabajadores no cualificados */.
    if (v=10) ocupac_2A =10 /* 10 Fuerzas armadas */.
  end if.
end repeat.
if (situac2_2A>=5 and situac2_2A<=11) ocupac_2A= -2 /* _NA No Aplicable - definido en la ECV sólo para ocupados*/.
if (gedad_2A=0) ocupac_2A=0 /* 0 ND (menor de 16 años) */.
execute.
Apply Dictionary

```

```

/ FROM BDH /SOURCE VARIABLES=ocupac_2A
/TARGET VARIABLES=ocupac_2A
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

lif (leff='EFF08') lthen.
*----- activestb1_2A: Actividad del Establecimiento de Trabajo del 2º adulto del Hogar (1ª Clasif)-----
String activestb1_2A (A5).
compute activestb1_2A=' '.
do repeat k=1 2 3 4 5 6 7 8 9
    / v = p6_4_1 p6_4_2 p6_4_3 p6_4_4 p6_4_5 p6_4_6 p6_4_7 p6_4_8 p6_4_9 .
do if (k=cod_2A) .
    if (v=1) activestb1_2A ='A+B' /* A+B Agricultura, ganadería, caza, silvicultura y pesca */.
    if (v=2 or v=3 or v=4) activestb1_2A ='C+D+E' /* C+D+E Indus. extractivas y manufact. Prod. y Distrib. de energía eléctrica, gas y agua */.
    if (v=5) activestb1_2A ='F' /* F Construcción */.
    if (v=6) activestb1_2A ='G' /* G Comercio; reparación de vehículos de motor, ciclomotores y artículos de uso doméstico */.
    if (v=7) activestb1_2A ='H' /* H Hostelería */.
    if (v=8) activestb1_2A ='I' /* I Transporte, almacenamiento y comunicaciones */.
    if (v=9) activestb1_2A ='J' /* J Intermediación financiera */.
    if (v=10) activestb1_2A ='K' /* K Actividades inmobiliarias y de alquiler; Servicios empresariales */.
    if (v=11) activestb1_2A ='L' /* L Administración pública y defensa; seguridad social obligatoria */.
    if (v=12) activestb1_2A ='M' /* M Educación */.
    if (v=13) activestb1_2A ='N' /* N Actividades sanitarias y veterinarias; servicios sociales */.
    if (v=14 or v=15 or v=16) activestb1_2A ='O+P+Q' /* O+P+Q Otras Activ y de servicios a la comunidad ; Serv personales. Hogares que emplea p.doméstico. Organ extraterritoriales */.
end if.
end repeat.
if (situac2_2A>=5 and situac2_2A<=11) activestb1_2A=' _NA' /* _NA No Aplicable - definido en la ECV sólo para ocupados */.
if (edad_2A=0) activestb1_2A=' _ND' /* _ND ND (menor de 16 años) */.
execute.
Apply Dictionary
/ FROM BDH /SOURCE VARIABLES=activestb1_2A
/TARGET VARIABLES=activestb1_2A
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
lifend.

lif (leff='EFF11') lthen.
*----- activestb2_2A: Actividad del Establecimiento de Trabajo del 2º Adulto del Hogar -----
String activestb2_2A (A3).
compute activestb2_2A=' '.
do repeat k=1 2 3 4 5 6 7 8 9
    / v = p6_4_1 p6_4_2 p6_4_3 p6_4_4 p6_4_5 p6_4_6 p6_4_7 p6_4_8 p6_4_9 .
do if (k=cod_2A) .
    if (v=1) activestb2_2A ='A' /* A Agricultura, ganadería, silvicultura y pesca */.
    if (v=2) activestb2_2A ='B' /* B Industrias extractivas */.
    if (v=3) activestb2_2A ='C' /* C Industria manufacturera */.
    if (v=4) activestb2_2A ='D' /* D Suministro de energía eléctrica, gas, vapor y aire ac ... */.
    if (v=5) activestb2_2A ='E' /* E Suministro de agua, actividades de saneamiento, .... */.
    if (v=6) activestb2_2A ='F' /* F Construcción */.
    if (v=7) activestb2_2A ='G' /* G Comercio al por mayor y al por menor; reparación ... */.
    if (v=8) activestb2_2A ='H' /* H Transporte y almacenamiento */.
    if (v=9) activestb2_2A ='I' /* I Hostelería */.
    if (v=10) activestb2_2A ='J' /* J Información y comunicaciones */.
    if (v=11) activestb2_2A ='K' /* K Actividades financieras y de seguros */.
    if (v=12) activestb2_2A ='L' /* L Actividades inmobiliarias */.
    if (v=13) activestb2_2A ='M' /* M Actividades profesionales, científicas y técnicas */.
    if (v=14) activestb2_2A ='N' /* N Actividades administrativas y servicios auxiliares */.
    if (v=15) activestb2_2A ='O' /* O Administración pública y defensa; seguridad social obligatoria */.
    if (v=16) activestb2_2A ='P' /* P Educación */.
    if (v=17) activestb2_2A ='Q' /* Q Actividades sanitarias y de servicios sociales */.
    if (v=18) activestb2_2A ='R' /* R Actividades artísticas, recreativas y de entretenimiento */.
    if (v=19) activestb2_2A ='S' /* S Otros servicios */.
    if (v=20) activestb2_2A ='T' /* T Actividades de los hogares como empleadores doméstico */.
    if (v=21) activestb2_2A ='U' /* U Actividades de organizaciones y organismos extraterritoriales */.
end if.
end repeat.
if (situac2_2A>=5 and situac2_2A<=11) activestb2_2A=' _NA' /* _NA No Aplicable - definido en la ECV sólo para ocupados */.
if (edad_2A=0) activestb2_2A=' _ND' /* _ND ND (menor de 16 años) */.
execute.
Apply Dictionary
/ FROM BDH /SOURCE VARIABLES=activestb2_2A
/TARGET VARIABLES=activestb2_2A
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
lifend.

lif (leff='EFF08') lthen.
*----- activestb1_2A_m: Actividad del establecimiento laboral (1ª clasif) -----
RECODE activestb1_2A (" ' '=SYSMIS) (_NA=-2) (_ND=0) (A+B=1) (C+D+E=2) (F=3) (G=4) (H=5) (I=6) (J=7)

```

```

('K'=8) ('L'=9) ('M'=10) ('N'=11) ('O+P+Q'=12) INTO activestb1_2A_m.
execute.
Apply Dictionary
/ FROM BDH /SOURCE VARIABLES=activestb1_2A_m
/TARGET VARIABLES=activestb1_2A_m
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
lifend.

lif (!eff='EFF11') !then.
*----- activestb2_2A_m: Actividad del establecimiento laboral (2º clasif) -----
RECODE activestb2_2A (' ' =SYSMIS) ('_NA'=-2) ('_ND'=0) ('A'=1) ('B'=2) ('C'=3) ('D'=4) ('E'=5) ('F'=6) ('G'=7)
('H'=8) ('I'=9) ('J'=10) ('K'=11) ('L'=12) ('M'=13) ('N'=14) ('O'=15) ('P'=16)
('Q'=17) ('R'=18) ('S'=19) ('T'=20) ('U'=21) INTO activestb2_2A_m.
execute.
APPLY DICTIONARY
/ FROM ArchivoH /SOURCE VARIABLES=activestb2_2A_m
/TARGET VARIABLES=activestb2_2A_m
/FILEINFO /VARINFO ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.
lifend.

*----- situprof_2A: Situación Profesional del 2º Adulto del Hogar -----
compute situprof_2A=$sysmis.
lif (!eff='EFF08') !then.
do repeat k=1 2 3 4 5 6
/ v1 = p4_103s1_1 p4_103s1_2 p4_103s1_3 p4_103s1_4 p4_103s1_5 p4_103s1_6
/ v2 = p4_103s2_1 p4_103s2_2 p4_103s2_3 p4_103s2_4 p4_103s2_5 p4_103s2_6
/ v3 = p4_103s3_1 p4_103s3_2 p4_103s3_3 p4_103s3_4 p4_103s3_5 p4_103s3_6
/ v4 = p4_103s4_1 p4_103s4_2 p4_103s4_3 p4_103s4_4 p4_103s4_5 p4_103s4_6
/ w1 = p4_109_1 p4_109_2 p4_109_3 p4_109_4 p4_109_5 p4_109_6 .
do if (situac2_2A=3 or situac2_2A=4) .
do if (k=1).
if ((v1=cod_2A or v2=cod_2A or v3=cod_2A or v4=cod_2A) and (w1 - 1 > 1)) situprof_2A=1 /* 1 Empleador
*/.
if ((v1=cod_2A or v2=cod_2A or v3=cod_2A or v4=cod_2A) and (w1 - 1 = 1)) situprof_2A = 2 /* 2 Empresario
sin asalariados o trabajador independiente */.
else if (k>1).
if ((v1=cod_2A or v2=cod_2A or v3=cod_2A or v4=cod_2A) and (w1 - 1 > 1)) situprof_2A=min(situprof_2A,1)
/* 1 Empleador */.
if ((v1=cod_2A or v2=cod_2A or v3=cod_2A or v4=cod_2A) and (w1 - 1 = 1)) situprof_2A=min(situprof_2A,2)
/* 2 Empresario sin asalariados o trabajador independiente */.
end if.
end if.
end repeat.
lifend
lif (!eff='EFF11') !then.
do repeat k=1 2 3 4 5 6
/ v1 = p4_103s1_1 p4_103s1_2 p4_103s1_3 p4_103s1_4 p4_103s1_5 p4_103s1_6
/ v2 = p4_103s2_1 p4_103s2_2 p4_103s2_3 p4_103s2_4 p4_103s2_5 p4_103s2_6
/ v3 = p4_103s3_1 p4_103s3_2 p4_103s3_3 p4_103s3_4 p4_103s3_5 p4_103s3_6
/ v4 = p4_103s4_1 p4_103s4_2 p4_103s4_3 p4_103s4_4 p4_103s4_5 p4_103s4_6
/ w1 = p4_109_1 p4_109_2 p4_109_3 p4_109_4 p4_109_5 p4_109_6
/ w2 = p4_109a_1 p4_109a_2 p4_109a_3 p4_109a_4 p4_109a_5 p4_109a_6 .
do if (situac2_2A=3 or situac2_2A=4) .
do if (k=1).
if ((v1=cod_2A or v2=cod_2A or v3=cod_2A or v4=cod_2A) and (w1 - 1*(w2=1)>1) ) situprof_2A=1 /* 1
Empleador */.
if ((v1=cod_2A or v2=cod_2A or v3=cod_2A or v4=cod_2A) and (w1 - 1*(w2=1)=1) ) situprof_2A = 2 /* 2
Empresario sin asalariados o trabajador independiente */.
else if (k>1).
if ((v1=cod_2A or v2=cod_2A or v3=cod_2A or v4=cod_2A) and (w1 - 1*(w2=1)>1) ) situprof_2A=min(situprof_2A,1)
/* 1 Empleador */.
if ((v1=cod_2A or v2=cod_2A or v3=cod_2A or v4=cod_2A) and (w1 - 1*(w2=1)=1) ) situprof_2A=min(situprof_2A,2)
/* 2 Empresario sin asalariados o trabajador independiente */.
end if.
end if.
end repeat.
lifend
if (situac2_2A=1 or situac2_2A=2) situprof_2A=3 /* 3 Asalariado */.
/* 4 Otra situación (Ayuda Familiar) */.
if (situac2_2A>=5 and situac2_2A<=11) situprof_2A=-2 /* -2 No Aplicable - definido en la ECV sólo para ocupados */.
if (gedad_2A=0) situprof_2A=0 /* 0 ND (menor de 16 años) */.
execute.

Apply Dictionary
/ FROM BDH /SOURCE VARIABLES=situprof_2A
/TARGET VARIABLES=situprof_2A
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*----- tipcont_2A: Tipo de contrato del 2º Adulto del Hogar -----
compute tipcont_2A=$sysmis.
do repeat k=1 2 3 4 5 6 7 8 9

```



```

        / v1 = p6_13_1_1 p6_13_2_1 p6_13_3_1 p6_13_4_1 p6_13_5_1 p6_13_6_1 p6_13_7_1 p6_13_8_1 p6_13_9_1
        / v2 = p6_13_1_2 p6_13_2_2 p6_13_3_2 p6_13_4_2 p6_13_5_2 p6_13_6_2 p6_13_7_2 p6_13_8_2 p6_13_9_2
        / v3 = p6_13_1_3 p6_13_2_3 p6_13_3_3 p6_13_4_3 p6_13_5_3 p6_13_6_3 p6_13_7_3 p6_13_8_3 p6_13_9_3 .
do if (k=cod_2A) .
    if (v1=2 or v2=2 or v3=2) tipcont_2A=2      /* 2 Contrato temporal de duración determinada */.
    if (v1=1 or v2=1 or v3=1) tipcont_2A=1      /* 1 Contrato fijo de duración indefinida */.
end if.
end repeat.
if (situac2_2A>=3 and situac2_2A<=11) tipcont_2A=-2      /* -2 No Aplicable - definido en la ECV sólo para asalariados-
ocupados*/.
if (gedad_2A=0) tipcont_2A=0                        /* 0 ND (menor de 16 años) */.
execute.

Apply Dictionary
/ FROM BDH /SOURCE VARIABLES=tipcont_2A
/TARGET VARIABLES=tipcont_2A
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

!if (!left='EFF08') !then.
*----- cse_2A: Categoría socio-económica del 2º Adulto del Hogar -----
compute cse_2A=$sysmis.
do if (situprof_2A =1).
    do if (activestb1_2A ='A+B' or ocupac_2A=6) .
        compute cse_2A=1      /* 1 Empres. Agrar. c/Asal */.
    else if (activestb1_2A ge 'C' and activestb1_2A le 'Q') .
        compute cse_2A=2      /* 2 Empres. No Agr. c/Asal */.
    else .
        compute cse_2A=3      /* 3 Otros Empresarios c/Asal No clasif */.
    end if.
else if (situprof_2A =2).
    do if (activestb1_2A ='A+B' or ocupac_2A=6) .
        compute cse_2A=4      /* 4 Empres. Agrar. s/Asal */.
    else if (activestb1_2A ge 'C' and activestb1_2A le 'Q') .
        compute cse_2A=5      /* 5 Empres. No Agr. s/Asal */.
    else .
        compute cse_2A=6      /* 6 Otros Empresarios s/Asal No clasif */.
    end if.
else if (situprof_2A =3).
    do if (activestb1_2A ='A+B' or ocupac_2A=6) .
        compute cse_2A=7      /* 7 Asalariados Agrarios */.
    else if (activestb1_2A ge 'C' and activestb1_2A le 'Q') .
        do if (ocupac_2A=1 or ocupac_2A=2).
            compute cse_2A=8      /* 8 Cuadros Super. No Agr. */.
        else if (ocupac_2A>=3 and ocupac_2A<=8).
            compute cse_2A=9      /* 9 Resto Empleados cualif. No Agr. */.
        else if (ocupac_2A=9) .
            compute cse_2A=10      /* 10 Empleados no cualif. No Agrar. */.
        else if (ocupac_2A=10) .
            compute cse_2A=11      /* 11 Fuerzas Armadas */.
        end if.
    else .
        compute cse_2A=12      /* 12 Otros Asalariados No clasif */.
    end if.
else if (situac2_2A =5) .
    compute cse_2A=13      /* 13 Parados */.
else if (situac2_2A =7).
    compute cse_2A=15      /* 15 Retirado */.
else if (situac2_2A =6).
    compute cse_2A=14      /* 14 Formación */.
else if (situac2_2A =8 or situac2_2A =10 or situac2_2A =11).
    compute cse_2A=16      /* 16 Otros Inactivos */.
end if.
if (gedad_2A=0) cse_2A=0      /* 0 ND (menor de 16 años) */.
!lifend.
!if (!left='EFF11') !then.
*----- cse_2A: Categoría socio-económica del 2º Adulto del Hogar -----
compute cse_2A=$sysmis.
do if (situprof_2A =1).
    do if (activestb2_2A ='A' or ocupac_2A=6) .
        compute cse_2A=1      /* 1 Empres. Agrar. c/Asal */.
    else if (activestb2_2A ge 'B' and activestb2_2A le 'U') .
        compute cse_2A=2      /* 2 Empres. No Agr. c/Asal */.
    else .
        compute cse_2A=3      /* 3 Otros Empresarios c/Asal No clasif */.
    end if.
else if (situprof_2A =2).
    do if (activestb2_2A ='A' or ocupac_2A=6) .
        compute cse_2A=4      /* 4 Empres. Agrar. s/Asal */.
    else if (activestb2_2A ge 'B' and activestb2_2A le 'U') .
        compute cse_2A=5      /* 5 Empres. No Agr. s/Asal */.
    else .

```

```

        compute cse_2A=6          /* 6 Otros Empresarios s/Asal No clasif */.
    end if.
else if (situac2_2A=3).
    do if (activestb2_2A='A' or ocupac_2A=6) .
        compute cse_2A=7          /* 7 Asalariados Agrarios */.
    else if (activestb2_2A ge 'B' and activestb2_2A le 'U') .
        do if (ocupac_2A=1 or ocupac_2A=2).
            compute cse_2A=8          /* 8 Cuadros Super. No Agr. */ .
        else if (ocupac_2A>=3 and ocupac_2A<=8).
            compute cse_2A=9          /* 9 Resto Empleados cualif. No Agr. */.
        else if (ocupac_2A=9) .
            compute cse_2A=10          /* 10 Empleados no cualif. No Agrar. */.
        else if (ocupac_2A=10) .
            compute cse_2A=11          /* 11 Fuerzas Armadas */.
        end if.
    else .
        compute cse_2A=12          /* 12 Otros Asalariados No clasif */.
    end if.
else if (situac2_2A=5) .
    compute cse_2A=13          /* 13 Parados */.
else if (situac2_2A=7).
    compute cse_2A=15          /* 15 Retirado */.
else if (situac2_2A=6).
    compute cse_2A=14          /* 14 Formación */ .
else if (situac2_2A=8 or situac2_2A=10 or situac2_2A=11).
    compute cse_2A=16          /* 16 Otros Inactivos */.
end if.
if (gedad_2A=0) cse_2A=0 /* 0 ND (menor de 16 años) */.
!lifend.
execute.
Apply Dictionary
/ FROM BDH /SOURCE VARIABLES=cse_2A
/TARGET VARIABLES=cse_2A
/FILEINFO /VARINFO VARLABEL ALIGNMENT FORMATS LEVEL MISSING VALLABELS=REPLACE WIDTH.

*=====,
* Otras .
*=====,

*----- riquezazet: Riqueza Neta. IIC para calcular la V de Cramer -----,
!Calcula_Grupos_Intercuantilicos segm=ano peso=w var=riquezazet ngrupos=10 pi=sí .

*----- Fuente: Fuente/Encuesta -----,

Numeric fuente (F2.0).
Variable Labels fuente 'Encuesta Fuente'.
!if (!left='EFF08') !then
Value Labels fuente 21 'EFF-08'.
compute fuente=21.
!lifend
!if (!left='EFF11') !then
Value Labels fuente 22 'EFF-11'.
compute fuente=22.
!lifend
execute.

*----- anoe: Año de referencia de la Encuesta -----,

Numeric anoe (F4.0).
Variable Labels anoe 'Año de referencia de la Encuesta'.
!let !etiqueta=!quote(!concat('año ',!ano))
Value Labels anoe !ano !quote(!etiqueta).
compute anoe=!ano.
execute.

*----- pais: País -----,

Numeric pais (F3.0).
Variable Labels pais 'País'.
Value Labels pais 34 'España'.
compute pais=34.
execute.

*----- wpc: Factor de elevación para ratios pc de los hogares -----,

Numeric wpc (F12.6).
Variable Labels wpc 'Factor de elevación para ratios pc de los hogares'.
compute wpc=w*nmh.
execute.

```


Anexo 2. Programa informático que realiza el *matching* entre los archivos que se fusionan

```
*****
* Proceso de matching.
*****
*-----
* Hay que especificar aquí los siguientes parámetros:
*-----
* nx=número de co-variables para el matching (en el maestro y en el de control).
* ny=número de variables auxiliares en el maestro (al menos 1, aunque sea nula).
* nz=número de variables matching en el archivo de control.
* seg=valor que identifica el segmento a fusionar
* ncasos= número de casos en el segmento del archivo maestro .
* n_ini= número del caso inicial del proceso de matching .
* n_fin= número del caso final del proceso de matching.
* nrpart= número máximo de casos de cada parte en las que se divide el proceso.
* id = variable de identificación de los casos en el maestro .
*-----
* NOTA: n_ini y n_fin se usan para determinar la parte del archivo a fusionar;
* pero dicha parte se fusiona completamente .
*-----

DEFINE !Proceso_de_matching
  (nx =!TOKENS(1)
   /ny =!TOKENS(1)
   /nz =!TOKENS(1)
   /seg =!TOKENS(1)
   /ncasos=!TOKENS(1)
   /n_ini=!TOKENS(1)
   /n_fin=!TOKENS(1)
   /nrpart=!TOKENS(1)
   /id =!TOKENS(1)
   /umbral=!TOKENS(1) !default(0.2) ).
!onexpand.

*--- Declara el directorio de ejecución del proceso de fusión.
File handle WkDir /NAME='UserDir\Work'.
CD WkDir.

*--- Abre el archivo de anotación de los tiempos de ejecución.
GET FILE='WkDir\_duracion_proceso.sav'.
DATASET NAME tiempos.
DATASET ACTIVATE tiempos.

*--- archivos de trabajo a realizar el proceso de matching -----
File handle pool_match_sav /NAME='WkDir\pool_para_match.sav'.
File handle casos_match_sav /NAME='WkDir\casos_match.sav'.

*//////////.

*----- homogeneiza ncasos y nrpart para poder compararlos numéricamente.
!!let !casos=!ncasos.
!do !i=1 !to 5.
  !if (!length(!casos)<5) !then
    !let !casos = !concat('0',!casos).
  !ifend.
!doend.

!!let !cmax=!nrpart.
!do !i=1 !to 5.
  !if (!length(!cmax)<5) !then
    !let !cmax = !concat('0',!cmax).
  !ifend.
!doend.

*----- cálculo de los intervalos de casos para las distintas partes.
!!let !limits0 = ".
!do !i=1 !to !casos !by !cmax.
  !!let !x=!.
  !do !j=1 !to 5.
    !if (!length(!x)<5) !then
      !let !x = !concat('0',!x).
    !ifend.
  !doend.
  !!let !limits0 = !concat(!limits0, ' ',!x, ' ').
!doend.
```

```

!!let !limits1 = ".
!do !i=!cmax !to !casos !by !cmax.
  !!let !x=!i.
  !do !j=1 !to 5.
    !if (!length(!x)<5) !then
      !!let !x = !concat('0',!x).
    !ifend.
  !doend.
  !!let !limits1 = !concat(!limits1,' ',!x,' ').
!doend.
!if (!x !ne !casos) !then
  !!let !limits1 = !concat(!limits1,' ',!casos,' ').
!ifend.

ECHO !quote(!concat("limites inferiores ',!limits0)).
ECHO !quote(!concat("limites superiores ',!limits1)).

*----- homogeneiza n_ini para poder compararlo numéricamente.
!!let !caso=!n_ini.
!do !i=1 !to 5.
  !if (!length(!caso)<5) !then
    !!let !caso = !concat('0',!caso).
  !ifend.
!doend.

*----- cálculo de la parte que contiene al caso n_ini.
!!let !lista0=!limits0.
!!let !lista1=!limits1.

!do !i=1 !to 99.
  !!let !lim0=!head(!lista0).
  !!let !lim1=!head(!lista1).
  !if (!caso >= !lim0 !and !caso <= !lim1) !then
    !!let !parte_ini = !i.
  !ifend.
  !!let !lista0=!tail(!lista0).
  !!let !lista1=!tail(!lista1).
!doend.

*----- homogeneiza n_fin para poder compararlo numéricamente.
!!let !caso = !n_fin.
!do !i=1 !to 5.
  !if (!length(!caso)<5) !then
    !!let !caso = !concat('0',!caso).
  !ifend.
!doend.

*----- cálculo de la parte que contiene al caso n_fin.
!!let !lista0=!limits0.
!!let !lista1=!limits1.

!do !i=1 !to 99.
  !!let !lim0=!head(!lista0).
  !!let !lim1=!head(!lista1).
  !if (!caso >= !lim0 !and !caso <= !lim1) !then
    !!let !parte_fin = !i.
  !ifend.
  !!let !lista0=!tail(!lista0).
  !!let !lista1=!tail(!lista1).
!doend.

*----- cálculo del número total de partes (n_partes).
!!let !lista0=!limits0.
!!let !lista1=!limits1.

!do !i=1 !to 99.
  !!let !lim0=!head(!lista0).
  !!let !lim1=!head(!lista1).
  !if (!casos >= !lim0 !and !casos <= !lim1) !then
    !!let !n_partes = !i.
  !ifend.
  !!let !lista0=!tail(!lista0).
  !!let !lista1=!tail(!lista1).
!doend.

*===== Inicio Bucle para las partes =====.

!do !n_parte= !parte_ini !to !parte_fin.

OUTPUT NEW NAME =salida_parte.
OUTPUT ACTIVATE salida_parte.

```

```

*calculo de los casos inicial c_ini y final c_fin de la parte

!!let !lista0=!limits0.
!!let !lista1=!limits1.

!do !i=1 !to !n_parte.
  !!let !c_ini=!head(!lista0).
  !!let !c_fin=!head(!lista1).
  !!let !lista0=!tail(!lista0).
  !!let !lista1=!tail(!lista1).
!doend.

!!let !mensaje = !quote(!concat('parte_ini=',!parte_ini,' parte_act=',!n_parte,' parte_fin=',!parte_fin,' n_partes=',!n_partes,'
caso_ini=',!c_ini,' caso_fin=',!c_fin,' ncasos=',!ncasos)).
ECHO !mensaje.

*--- crea vacío el directorio de trabajo para la Parte.
!!let !cmd1=!quote(!concat('MD Parte',!n_parte)).
HOST COMMAND=[!cmd1].

* define nuevos archivos de trabajo durante el proceso de matching -----.
!!let !directorio=!quote(!concat('WkDir\Parte',!n_parte)).
File handle ProcDir /NAME=!directorio.
* File handle nuevo_match_sav /NAME='ProcDir\nuevo_match.sav'.
File handle pool_parte_sav /NAME='ProcDir\pool_para_match.sav'.
File handle casos_parte_sav /NAME='ProcDir\casos_match.sav'.

*--- carga del archivo de trabajo <pool_parte> en el directorio de la parte ProcDir.
GET FILE='pool_match.sav'.
DATASET NAME pool_match.
DATASET ACTIVATE pool_match.
DATASET COPY pool_parte WINDOWS=HIDDEN.
DATASET ACTIVATE pool_parte.
select if ((!c_ini<=idx and idx<=!c_fin) or (source=0)).
save outfile='pool_parte.sav' /compressed.
DATASET CLOSE pool_match.

*--- Creación del archivo vacío <casos_parte> en el directorio de la parte ProcDir.
NEW FILE.
INPUT PROGRAM.
compute idx=$sysmis.
END FILE.
END INPUT PROGRAM.
select if (1=0).
execute.
DATASET COPY casos_parte WINDOWS=HIDDEN.
DATASET ACTIVATE casos_parte.
save OUTFILE='casos_parte.sav'.

*--- Escribe el tiempo de inicio del proceso de matching.
!!let !parte=!concat(Parte,!n_parte).
DATASET ACTIVATE tiempos.
DO IF ($casenum=!n_parte).
  COMPUTE Parte=!quote(!parte).
  COMPUTE T_ini_match=$time.
END IF.
EXECUTE.
save outfile='WkDir\_duracion_proceso.sav' /compressed.

*=====
*--- Proceso de matching de la parte seleccionada del archivo maestro.

DATASET CLOSE casos_parte.
DATASET CLOSE pool_parte.

!!let !Xvars=!concat('X',!nx).
!if (!nx>1) !then
  !!let !Xvars=!concat('X1 to ',!Xvars).
!ifend.
!!let !Yvars=!concat('Y',!ny).
!if (!ny>1) !then
  !!let !Yvars=!concat('Y1 to ',!Yvars).
!ifend.
!!let !Zvars=!concat('Z',!nz).
!if (!nz>1) !then
  !!let !Zvars=!concat('Z1 to ',!Zvars).
!ifend.

echo 'Inicio del emparejamiento de casos'.

```

```
!!let !peso=!concat('Peso_',!seg).
MATRIX.
```

* Lectura de la variable de análisis y la de ponderación.

*-----.

```
GET Vidm
/ File = 'ProcDir\pool_para_match.sav'
/ VARIABLES = idm
/ NAMES = Vidm_N
/ MISSING = ACCEPT
/ SYSMIS = -999.
```

```
GET XM
/ File = 'ProcDir\pool_para_match.sav'
/ VARIABLES = !Xvars
/ NAMES = XM_N
/ MISSING = ACCEPT
/ SYSMIS = -999.
```

```
GET YM
/ File = 'ProcDir\pool_para_match.sav'
/ VARIABLES = !Yvars
/ NAMES = YM_N
/ MISSING = ACCEPT
/ SYSMIS = -999.
```

```
GET Vidc
/ File = 'ProcDir\pool_para_match.sav'
/ VARIABLES = idc
/ NAMES = Vidc_N
/ MISSING = ACCEPT
/ SYSMIS = -999.
```

```
GET ZC
/ File = 'ProcDir\pool_para_match.sav'
/ VARIABLES = !Zvars
/ NAMES = ZC_N
/ MISSING = ACCEPT
/ SYSMIS = -999.
```

```
GET cod_ND
/ File = 'WkDir\Declarar_VariablesX_Matching.sav'
/ VARIABLES = cod_ND
/ NAMES = cod_ND_N
/ MISSING = ACCEPT
/ SYSMIS = -999.
```

```
GET EscX
/ File = 'WkDir\Declarar_VariablesX_Matching.sav'
/ VARIABLES = EscX
/ NAMES = EscX_N
/ MISSING = ACCEPT
/ SYSMIS = -999.
```

```
GET Peso
/ File = 'WkDir\Declarar_VariablesX_Matching.sav'
/ VARIABLES = !Peso
/ NAMES = Peso_N
/ MISSING = ACCEPT
/ SYSMIS = -999.
```

```
GET delta
/ File = 'WkDir\Declarar_VariablesX_Matching.sav'
/ VARIABLES = delta
/ NAMES = delta_N
/ MISSING = ACCEPT
/ SYSMIS = -999.
```

* Cálculo de estadísticos de las variables Xk.

*-----.

```
compute nX = NCOL(XM).
compute n = NROW(XM).
compute MaxX = CMAX( XM &* (XM ne -999) + -1.e20 &* (XM eq -999) ).
compute MinX = CMIN( XM &* (XM ne -999) + 1.e20 &* (XM eq -999) ).
compute nZ = NCOL(ZC).
```

* Extracción de casos del Archivo Maestro y de Control.

*-----.

```
compute nM = MAKE(1,n,1)*(Vidm>0).
compute nC = n-nM.
compute nM1 = nM+1.
compute XC = XM(nM1:n, : ) .
compute ZC = ZC(nM1:n, : ) .
compute Vidc = Vidc(nM1:n).
compute XM = XM(1:nM, : ) .
compute Vidm = Vidm(1:nM).
```

* Preparación de variables para la Fusión .

*-----.

```

compute nciclos=9 /* número de ciclos de búsqueda de parejas */ .
compute idciclo = MAKE(nM,1,-999) /* identificador del ciclo de búsqueda de parejas en el archivo de control */.
compute WM=YM(1:nM,1) /* pesos que quedan por asignar al principio de cada ciclo de búsqueda */ .
compute WC=ZC( , 1) /* pesos disponibles para asignación en los casos de control al principio de cada
ciclo de búsqueda */.

*--- si la suma de pesos del archivo de control es menor que en el maestro, se reescalan para igualar las sumas (pesos
disponibles) .
do if ( CSUM(WC) lt CSUM(WM) ).
  compute WC=WC &* CSUM(WM) &/ CSUM(WC).
end if.

* Preparación de matrices de índices, pesos y distancias de enlace.
*-----..

compute VNames= {Vidm_N, 'partic', 'idcmatch', 'wmatch', 'prox', XM_N, ZC_N}.

print nciclos.
LOOP #j=1 TO nciclos.
  * print #j.
  compute idciclo = MAKE(nM, 1, #j).
  compute idcmatch = MAKE(nM,1,-999) /* identificador de la pareja encontrada en el archivo de control */.
  compute prox = MAKE(nM,1,-999) /* proximidad con la pareja encontrada en el archivo de control */.
  compute wmatch = MAKE(nM,1,-999) /* peso correspondiente a la pareja encontrada en el archivo de control */.

  compute proxC = MAKE(nC,1,-999).
  compute ZM = MAKE(nM,nZ,-999).
  compute X_C = MAKE(nM,nX,-999).

  LOOP #i=1 TO nM.
  DO IF (WM(#i)>1.e-6).

    * Cálculo de distancias entre cada caso del M y los casos del C .
    *-----..
    compute Xki = make(nC,1,1)*XM(#i , : ) .
    compute Xkj = XC .
    compute NDk = make(nC,1,1)*T(cod_ND).
    compute XkiMiss = (Xki eq -999) or (Xki eq NDk) .
    compute XkjMiss = (XC eq -999) or (XC eq NDk) .
    compute Wkij = ( make(nC,1,1)*T(peso) ) &*(XkiMiss=0).
    compute Wkij = Wkij &/ (Rsum(Wkij)*make(1,nX,1)) .
    compute Dkij = ( make(nC,1,1)*T( (EscX=1) ) ) &* (1 - (Xki = Xkj)) &*(XkiMiss=0 and XkjMiss=0) +
(make(nC,1,1)*T(delta))&*(XkiMiss=1 or XkjMiss=1).
    compute Dkij = Dkij + (make(nC,1,1)*T((EscX=2 or EscX=3))) &* (abs(Xki - Xkj)/(make(nC,1,1)*( MaxX=MinX)+MaxX-
MinX)))&*(XkiMiss=0 and XkjMiss=0) + (make(nC,1,1)*T(delta))&*(XkiMiss=1 or XkjMiss=1).
    compute proxC= ( Wkij &* Dkij )*make(nX,1,1) + (9999 &* (WC=0)) /* si no hay peso disponible, la distancia se pone a 9999
(máximo artificial) */.

    * Enlace del caso-match y carga de sus características.
    *-----..
    compute numcaso={1:nC}* (GRADE(proxC)=1).
    compute proximid=proxC(numcaso).
    DO IF (proximid le !umbral or #j=1).
      compute prox(#i)=proxC(numcaso).
      compute idcmatch(#i)=Vidc(numcaso).
      compute wmatch(#i)= WM(#i)&*(WM(#i) le WC(numcaso)) + WC(numcaso)&*(WM(#i) gt WC(numcaso)).
      compute WM(#i) = WM(#i) - wmatch(#i).
      compute WC(numcaso) = WC(numcaso) - wmatch(#i).
      compute ZM(#i,:)=ZC(numcaso,:).
      compute X_C(#i,:)=XC(numcaso,:).
    END IF.

  END IF.
  END LOOP.

  DO IF (#j=1).
    SAVE {Vidm, idciclo, idcmatch, wmatch, prox, X_C, ZM }
    / OUTFILE = 'ProcDir\casos_match.sav'
    / NAMES = VNames .
  ELSE .
    SAVE {Vidm, idciclo, idcmatch, wmatch, prox, X_C, ZM }
    / NAMES = VNames .
  END IF.

  END LOOP.

  *-----..
  * Grabación del archivo casos_match.sav .
  *-----..
  * save {Xki, Xkj, Wkij, Dkij} /outfile='pruebas.sav'.

END MATRIX.

```



```

* GET FILE='pruebas.sav'.
* DATASET NAME pruebas.
* dataset activate pruebas.

echo 'fin del emparejamiento de casos'.

*===== Reestructuración de los casos emparejados =====.
GET FILE='ProcDir\casos_match.sav'.
DATASET NAME casos.
DATASET ACTIVATE casos.
RECODE partic to prox !Zvars (-999=SYSMIS).
select if (not missing (idcmatch)).
EXECUTE.
sort cases by idm partic prox.
Formats partic (F2.0).

save outfile='ProcDir\casos_match.sav' / compressed.
dataset close casos.

*=====

*--- Escribe el tiempo de finalización del proceso de matching.
DATASET ACTIVATE tiempos.
DO IF ($casenum=ln_parte).
  COMPUTE T_fin_match=$time.
  COMPUTE T_matching=T_fin_match-T_ini_match.
END IF.
EXECUTE.
save outfile='WkDir\_duracion_proceso.sav'/compressed.

* DATASET CLOSE casos_parte.
* DATASET CLOSE pool_parte.
ERASE FILE='ProcDir\pool_para_match.sav'.

OUTPUT SAVE NAME=salida_parte
  OUTFILE='ProcDir\_Proceso_de_matching.spv' LOCK=NO.
OUTPUT CLOSE salida_parte.

!doend.
*===== Fin Bucle para las partes =====.

NEW FILE.
DATASET CLOSE ALL.

!ENDDEFINE.

```

Anexo 3. Programa de cálculo de los indicios de incongruencia

```
*****.
* 040d_Cálculo_de_discordancias.sps
*****.

DEFINE !Calculo_de_discordancias (
    Vseg = !TOKENS (1)
    / idh = !TOKENS (1)
    / filtro = !TOKENS (1)
    / Yvar = !TOKENS (1)
    / Xvar = !TOKENS (1)
    / Rel = !TOKENS (1) !DEFAULT (Pos)
    / anno = !TOKENS (1)
    / Corr = !TOKENS (1) !DEFAULT (Ocu)
    / TipInd = !TOKENS (1) !DEFAULT (INC)
    / NumInd = !TOKENS (1) !DEFAULT (000)
    / Umbral = !TOKENS (1) !DEFAULT ("0" ) .

!onexpand.

*=====
*--- Proceso de cálculo de concordancias .
!!let !INCONG = !concat(!TipInd,'_I',!NumInd)
!!let !JNCONG = !concat(!TipInd,'_J',!NumInd)
!!let !DNCONG = !concat(!TipInd,'_D',!NumInd)
!!let !FNCONG = !concat(!TipInd,'_',!NumInd)
!!let !Vmbrai = !Unquote(!Umbral)

New file.
dataset name discordancias.

DATASET ACTIVATE BDH.
DATASET COPY Aux.
DATASET ACTIVATE Aux.
select if (ano = !anno and !filtro = 1).
execute.

MATRIX.

* Lectura de la variable de análisis y la de ponderación.
*-----
GET Datos
/ File = *
/ VARIABLES = !Vseg !idh !Yvar !Xvar
/ NAMES = Datos_N
/ MISSING = OMIT.

compute ini = make(50,1,-999) .
compute fin = make(50,1,-999) .

compute n = NROW(Datos).
LOOP #i=1 TO n.
DO IF (#i=1).
    compute nseg = 1 .
    compute ini(nseg) = 1 .
    compute #segm = Datos(#i,1) .
ELSE IF (Datos(#i,1) ne #segm) .
    compute fin(nseg) = #i -1 .
    compute nseg = nseg + 1 .
    compute ini(nseg) = #i .
    compute #segm = Datos(#i,1) .
END IF.
END LOOP.
compute fin(nseg) = n .
*-----

LOOP #j=1 TO nseg.

    compute nX = fin(#j) - ini(#j) +1.

Do if (nX>1).

    compute IOAN = MAKE(nX,1,-999) .
    compute IOAP = MAKE(nX,1,-999) .
    compute ISAN = MAKE(nX,1,-999) .
    compute ISAP = MAKE(nX,1,-999) .
    compute JOAN = MAKE(nX,1,-999) .
```

```

compute JOAP = MAKE(nX,1,-999) .
compute JSAN = MAKE(nX,1,-999) .
compute JSAP = MAKE(nX,1,-999) .
compute DOAN = MAKE(nX,1,-999) .
compute DOAP = MAKE(nX,1,-999) .
compute DSAN = MAKE(nX,1,-999) .
compute DSAP = MAKE(nX,1,-999) .

```

```

compute Y = Datos(ini(#):fin(#),3) .
compute X = Datos(ini(#):fin(#),4) .

```

```

LOOP #i=1 TO nX.
  compute Xi = make(nX,1,X(#i)) .
  compute Yi = make(nX,1,Y(#i)) .
  compute IOAN(#i) = T(Xi < X) * (Yi <= Y).
  compute IOAP(#i) = T(Xi > X) * (Yi <= Y).
  compute ISAN(#i) = T(Xi > X) * (Yi >= Y).
  compute ISAP(#i) = T(Xi < X) * (Yi >= Y).
  compute JOAN(#i) = ABS(T((Xi - X)&*(Xi<X)) * ((Yi - Y)&*(Yi<=Y))).
  compute JOAP(#i) = ABS(T((Xi - X)&*(Xi>X)) * ((Yi - Y)&*(Yi<=Y))).
  compute JSAN(#i) = ABS(T((Xi - X)&*(Xi>X)) * ((Yi - Y)&*(Yi>=Y))).
  compute JSAP(#i) = ABS(T((Xi - X)&*(Xi<X)) * ((Yi - Y)&*(Yi>=Y))).
  compute DOAN(#i) = T( (Xi - X)&**2 + (Yi - Y)&**2 ) * ((Xi<X)&*(Yi<=Y)) .
  compute DOAP(#i) = T( (Xi - X)&**2 + (Yi - Y)&**2 ) * ((Xi>X)&*(Yi<=Y)) .
  compute DSAN(#i) = T( (Xi - X)&**2 + (Yi - Y)&**2 ) * ((Xi>X)&*(Yi>=Y)) .
  compute DSAP(#i) = T( (Xi - X)&**2 + (Yi - Y)&**2 ) * ((Xi<X)&*(Yi>=Y)) .
END LOOP.

```

```

DO IF (!quote(!Corr)='Ocu').
  DO IF (!quote(!Rel)='Pos').
    compute !INCONG = IOAP .
    compute !JNCONG = JOAP .
    compute !DNCONG = DOAP .
  ELSE IF (!quote(!Rel)='Neg').
    compute !INCONG = IOAN .
    compute !JNCONG = JOAN .
    compute !DNCONG = DOAN .
  END IF.
ELSE IF (!quote(!Corr)='Sob').
  DO IF (!quote(!Rel)='Pos').
    compute !INCONG = ISAP .
    compute !JNCONG = JSAP .
    compute !DNCONG = DSAP .
  ELSE IF (!quote(!Rel)='Neg').
    compute !INCONG = ISAN .
    compute !JNCONG = JSAN .
    compute !DNCONG = DSAN .
  END IF.
END IF.

```

```

compute INC = !INCONG / (nX-1) .
compute MINC = T(!INCONG) * make(nX,1,1/(nX*(nX-1))) &* make(nX,1,1) .
compute DTINC = SQRT(T(!INCONG) * !INCONG &* make(nX,1,1/(nX*(nX-1)&**2)) - MINC&**2).
compute DTINC = DTINC &* (DTINC>0) + 1.e-6 &* (DTINC=0) .
compute ZINC = (INC-MINC) &/ DTINC.
* compute !INCONG = ZINC .
* compute !INCONG = (ZINC - make(nX,1,!Vmbral)) &* (ZINC > make(nX,1,!Vmbral)) .

```

```

compute JNC = !JNCONG / (nX-1) .
compute MJNC = T(!JNCONG) * make(nX,1,1/(nX*(nX-1))) &* make(nX,1,1) .
compute DTJNC = SQRT(T(!JNCONG) * !JNCONG &* make(nX,1,1/(nX*(nX-1)&**2)) - MJNC&**2).
compute DTJNC = DTJNC &* (DTJNC>0) + 1.e-6 &* (DTJNC=0) .
compute ZJNC = (JNC-MJNC) &/ DTJNC .
* compute !JNCONG = ZJNC .
* compute !JNCONG = (ZJNC - make(nX,1,!Vmbral)) &* (ZJNC > make(nX,1,!Vmbral)) .

```

```

compute DNC = !DNCONG / (nX-1) .
compute MDNC = T(!DNCONG) * make(nX,1,1/(nX*(nX-1))) &* make(nX,1,1) .
compute DTDNC = SQRT(T(!DNCONG) * !DNCONG &* make(nX,1,1/(nX*(nX-1)&**2)) - MDNC&**2).
compute DTDNC = DTDNC &* (DTDNC>0) + 1.e-6 &* (DTDNC=0) .
compute ZDNC = (DNC-MDNC) &/ DTDNC .
* compute !DNCONG = ZDNC .
* compute !DNCONG = (ZDNC - make(nX,1,!Vmbral)) &* (ZDNC > make(nX,1,!Vmbral)) .

```

else.

```

compute INC = 0 .
compute JNC = 0 .
compute DNC = 0 .

```

end if.

```

DO IF (#j=1).
*   compute VNames= {Datos_N(1:2) , 'IOAN', 'IOAP', 'ISAN', 'ISAP', 'INC', 'MINC', 'DTINC', 'ZINC', !quote(!INCONG), 'JOAN',
'JOAP', 'JSAN', 'JSAP', 'JNC', 'MJNC', 'DTJNC', 'ZJNC', !quote(!JNCONG) }.
*   SAVE { Datos(ini(#j):fin(#j),1:2), IOAN, IOAP, ISAN, ISAP, INC, MINC, DTINC, ZINC, !INCONG, JOAN, JOAP,
JSAN, JSAP, JNC, MJNC, DTJNC, ZJNC, !JNCONG }.
compute VNames= {Datos_N(1:4) , !quote(!INCONG), !quote(!JNCONG), !quote(!DNCONG) }.
SAVE { Datos(ini(#j):fin(#j),1:4), INC, JNC, DNC }
/ OUTFILE = discordancias
/ NAMES = VNames .
ELSE .
*   SAVE { Datos(ini(#j):fin(#j),1:4), IOAN, IOAP, ISAN, ISAP, INC, MINC, DTINC, ZINC, !INCONG, JOAN, JOAP,
JSAN, JSAP, JNC, MJNC, DTJNC, ZJNC, !JNCONG }.
SAVE { Datos(ini(#j):fin(#j),1:4), INC, JNC, DNC }
/ NAMES = VNames .
END IF.

END LOOP.

END MATRIX.
DATASET close Aux.

*===== Reestructuración de los casos emparejados =====.

DATASET ACTIVATE discordancias.
FACTOR VARIABLES !INCONG !JNCONG !DNCONG
/CRITERIA FACTORS(1)
/ROTATION NOROTATE
/SAVE REG(ALL).

compute FAC1_1 = FAC1_1 .
if (missing (FAC1_1)) FAC1_1 = (!INCONG + !JNCONG + !DNCONG) / 3 .
compute FAC1_1 = FAC1_1 * (FAC1_1>Vmbra1) + !Vmbra1 * (FAC1_1<=Vmbra1).
execute.
delete variables !Yvar !Xvar !INCONG !JNCONG !DNCONG.
rename variables (FAC1_1 = !FNCONG).
variable label !FNCONG !quote(!FNCONG).

DATASET ACTIVATE BDH.
!if (!anno = 2008) !then.
MATCH FILES FILE=*
/TABLE=discordancias
/BY= !Vseg !idh.
execute.
!else.
UPDATE FILE=*
/FILE=discordancias
/BY= !Vseg !idh.
execute.

DATASET close discordancias.

IENDDEFINE.

```

Anexo 4. Programa que calcula "k" para una partida y la corrige

*-----.

* Obtiene la corrección "O_x" de la variable "x0" para que, ponderada mediante "pesox", redistribuya la ocultación "Ox" sólo entre los casos seleccionados en la variable "filtro".

* la corrección se hace mediante una tasa de ocultación lineal, "t", función de una covariable explicativa "y0".

* La ocultación asignada a un caso de valores (x0,y0) será $t \cdot x$, siendo; $t = a + b \cdot y$, $x = x0 - \min X$, $y = y0 - \min Y$.

* y la pendiente "b" puede fijarse en una fracción "k" de la máxima pendiente que permitiría redistribuir las ocultaciones por este procedimiento.

* La covariable "y0" se puede obtener a partir de una o dos variables explicativas "x1" y "x2" que deben estar ordenadas en el mismo sentido con relación a la magnitud de la ocultación a imputar.

* En el caso de usar dos variables explicativas, el valor de la covariable para cada caso se obtiene como la distancia euclídea de sus coordenadas estandarizadas en el cuadrado $[0;1] \times [0;1]$ al vertice origen.

*-----.

```

DEFINE !calcula_correccion_optima_tasa_lineal (
    partida = !TOKENS(1)
    / Ox      = !TOKENS(1)
    / filtro  = !TOKENS(1)
    / x0      = !TOKENS(1)
    / pesox   = !TOKENS(1)
    / k       = !TOKENS(1) !default(0)
    / x1      = !Enclose('(',')' ) .

    !let !kvar = !concat('K_', !partida)
    !let !Oh = !concat('Oh_', !partida)

    USE ALL.
    compute O_x = 0          /* O_x será la corrección obtenida */ .
    compute y0 = 0          /* y0 será la norma del vector de intensidad de la corrección */.
    compute vfiltro = !unquote(!filtro) .
    FILTER BY vfiltro.
    EXECUTE.

```

*--- cálculo de la covariable (para la tasa de ocultación lineal) -----.

```

!DO !vx !IN (!x1)

AGGREGATE
/OUTFILE=* MODE=ADDVARIABLES OVERWRITE=yes
/PRESORTED
/BREAK=ano
/x1min=MIN(!vx )
/x1max=MAX(!vx ).

if (x1max>x1min) y0=y0 + ((!vx-x1min)/(x1max-x1min))**2.
execute.

!DOEND
compute y0=SQRT(y0).
execute.

```

*--- cálculo de variables auxiliares para el cálculo de la ocultación individual.

```

WEIGHT BY !pesox.
AGGREGATE
/OUTFILE=* MODE=ADDVARIABLES OVERWRITE=yes
/PRESORTED
/BREAK=ano
/y0min=MIN( y0)
/x0min=MIN(!x0)
/y0max=MAX( y0)
/x0max=MAX(!x0).
WEIGHT OFF.

```

```
compute y= (y0-y0min)**(1/1).
compute x=(!x0-x0min)**(1/1).
compute xy=y*x.
execute.
```

```
WEIGHT BY !pesox.
AGGREGATE
/OUTFILE=* MODE=ADDVARIABLES OVERWRITE=yes
/PRESORTED
/BREAK=ano
/Sx=SUM(x)
/Sxy=SUM(xy)
/Sw=N(y) .
WEIGHT OFF.
```

*--- cálculo variables auxiliares y cálculo óptimo de k -----.

```
!if (!k !eq 90) !then
WEIGHT BY !pesox.
AGGREGATE
/OUTFILE=* MODE=ADDVARIABLES OVERWRITE=yes
/PRESORTED
/BREAK=ano idh_bdh
/Shx=SUM(x)
/Shxy=SUM(xy) .
WEIGHT OFF.
```

```
compute Oh = !Oh * (!Oh*!Ox>=0).
compute aaa = Shxy/Sxy - Shx/Sx.
compute bbb = aaa * (Oh - !Ox*Shx/Sx).
compute ccc = aaa**2.
execute.
```

```
WEIGHT BY !pesox.
AGGREGATE
/OUTFILE=* MODE=ADDVARIABLES OVERWRITE=yes
/PRESORTED
/BREAK=ano
/Sbbb=SUM(bbb)
/Sccc=SUM(ccc) .
WEIGHT OFF.
```

```
if (vfiltro=1) !kvar = max(0, min(1, Sbbb / !Ox / Sccc)) .
execute.
!lifend.
```

```
!if (!k !eq 99) !then
GET FILE='BDRenCb04Dir\BDH\Valores_optimos_k.sav'.
DATASET NAME aux WINDOW=FRONT.
DATASET ACTIVATE aux.
compute aaa=!kvar.
execute.
delete variables K_AIVPP_P to K_D_75_e_Bdh.
rename variables (aaa=!kvar).
```

```
DATASET ACTIVATE BDH.
MATCH FILES /FILE=*
/TABLE='aux'
/BY ano.
EXECUTE.
dataset close aux.
!lifend.
```

```
!if (!k !ne 90 !and !k !ne 99) !then
if (vfiltro=1) !kvar = !k.
execute.
!lifend.
```

*--- cálculo de la ocultación -----.

```

do if (x0max-x0min=0) .
  compute O_x = !Ox/Sw .
else if (y0max-y0min=0) .
  compute O_x = !Ox*x/Sx .
else .
  compute O_x = !Ox*x*((1-!kvar)/Sx + !kvar*y/Sxy) .
end if.
execute.

```

```

FILTER OFF.
USE ALL.
EXECUTE.

```

*--- actualización de variables corregidas -----.

```

compute !partida = !partida + O_x .
execute.
!if (!k !eq 90) !then
  delete variables O_x y0 to Sccc.
!else
  delete variables O_x y0 to Sw.
!lifend

```

```

!ENDDEFINE.

```

Anexo 5. Programa de corrección de la BDH

```
*=====,
* Proy: Tesis de Cristina Romero.
* SubProy: Creación y Mantenimiento de la Base de Datos Homogénea de Rentas.
* App: 040e_Corrig-ocultaciones .
*=====,
*=====,
*==>>> 040e_Corrig-ocultaciones.sps
*=====,

NEW FILE.
DATASET CLOSE ALL.

OUTPUT NEW NAME =salida.
OUTPUT ACTIVATE salida.

CD 'WorkDir'.

*=====,
* Cálculo de Ocultaciones agregadas desde la BDN.
*=====,

Get file='BDNDir\BDN.sav' / KEEP=pais to um
  Npers_Cne K_1_Cne B_2n_Cne B_2b_Cne B_3n_Cne B_3b_Cne B_2n3n_Cne B_2b3b_Cne
  D_11_r_Cne D_121_r_Cne D_122_r_Cne D_12_r_Cne D_1_r_Cne
  D_41_r_Cne D_41_e_Cne D_42_r_Cne D_4142_re_Cne D_44_r_Cne D_45_r_Cne D_45_e_Cne D_45_re_Cne D_4_r_Cne
D_4_e_Cne D_4_re_Cne B_5n_Cne B_5b_Cne
  D_51_e_Cne D_59_e_Cne D_5_e_Cne
  D_611_e_Cne D_612_r_Cne D_612_e_Cne D_612_er_Cne D_61_r_Cne D_61_e_Cne D_61_er_Cne
  D_621_r_Cne D_622_r_Cne D_623_r_Cne D_623_e_Cne D_624_r_Cne D_62_r_Cne D_62_e_Cne D_62_re_Cne
  D_71_e_Cne D_72_r_Cne D_75_r_Cne D_75_e_Cne D_7_r_Cne D_7_e_Cne B_6b_Cne B_6n_Cne
  AIVPP_P_Est AIVPP_CL_IPPH_Est AIVPP_CL_Seg_Est AIVPP_CL_TCE_Est AIVPP_OIN_Est CCFFEE_Est
  IMALQPRO_B_Est CCFRM_Est ORNC1_B_Est CSIMEMP_Est ORI_B_Est IMALQTER_B_Est CSEFHOG_Est
  CSIMr_Est ORNC2_B_Est ORNC3_B_Est INDNOVIDA1_B_Est ORNC4_B_Est INDNOVIDA2_B_Est
  OTRDOH_B_Est PRIMNOVIDA_B_Est OTPAOH_B_Est
  Npers_Nac K_1_Nac B_2n_Nac B_2b_Nac B_3n_Nac B_3b_Nac B_2n3n_Nac B_2b3b_Nac
  D_11_r_Nac D_121_r_Nac D_122_r_Nac D_12_r_Nac D_1_r_Nac
  D_41_r_Nac D_41_e_Nac D_42_r_Nac D_4142_re_Nac D_44_r_Nac D_45_r_Nac D_45_e_Nac D_45_re_Nac D_4_r_Nac
D_4_e_Nac D_4_re_Nac B_5n_Nac B_5b_Nac
  D_51_e_Nac D_59_e_Nac D_5_e_Nac
  D_611_e_Nac D_612_r_Nac D_612_e_Nac D_612_er_Nac D_61_r_Nac D_61_e_Nac D_61_er_Nac
  D_623_e_Nac D_62_r_Nac D_62_e_Nac D_62_re_Nac
  D_71_e_Nac D_72_r_Nac D_75_r_Nac D_75_e_Nac D_7_r_Nac D_7_e_Nac B_6b_Nac B_6n_Nac
  AIVPP_P AIVPP_CL_IPPH AIVPP_CL_Seg AIVPP_CL_TCE AIVPP_OIN CCFFEE
  IMALQPRO_B CCFRM ORNC1_B CSIMEMP ORI_B IMALQTER_B CSEFHOG
  CSIMr ORNC2_B ORNC3_B INDNOVIDA1_B ORNC4_B INDNOVIDA2_B
  OTRDOH_B PRIMNOVIDA_B OTPAOH_B .
DATASET NAME BDN.
DATASET ACTIVATE BDN.
select if ((ano=2008 or ano=2011) and region=90).
execute.

do repeat v1 =
  Npers_Cne K_1_Cne B_2n_Cne B_2b_Cne B_3n_Cne B_3b_Cne B_2n3n_Cne B_2b3b_Cne
  D_11_r_Cne D_121_r_Cne D_122_r_Cne D_12_r_Cne D_1_r_Cne
  D_41_r_Cne D_41_e_Cne D_42_r_Cne D_4142_re_Cne D_44_r_Cne D_45_r_Cne D_45_e_Cne D_45_re_Cne D_4_r_Cne
D_4_e_Cne D_4_re_Cne B_5n_Cne B_5b_Cne
  D_51_e_Cne D_59_e_Cne D_5_e_Cne
  D_611_e_Cne D_612_r_Cne D_612_e_Cne D_612_er_Cne D_61_r_Cne D_61_e_Cne D_61_er_Cne
  D_623_e_Cne D_62_r_Cne D_62_e_Cne D_62_re_Cne
  D_71_e_Cne D_72_r_Cne D_75_r_Cne D_75_e_Cne D_7_r_Cne D_7_e_Cne B_6b_Cne B_6n_Cne
  AIVPP_P_Est AIVPP_CL_IPPH_Est AIVPP_CL_Seg_Est AIVPP_CL_TCE_Est AIVPP_OIN_Est CCFFEE_Est
  IMALQPRO_B_Est CCFRM_Est ORNC1_B_Est CSIMEMP_Est ORI_B_Est IMALQTER_B_Est CSEFHOG_Est
  CSIMr_Est ORNC2_B_Est ORNC3_B_Est INDNOVIDA1_B_Est ORNC4_B_Est INDNOVIDA2_B_Est
  OTRDOH_B_Est PRIMNOVIDA_B_Est OTPAOH_B_Est
  / v2 =
  Npers_Nac K_1_Nac B_2n_Nac B_2b_Nac B_3n_Nac B_3b_Nac B_2n3n_Nac B_2b3b_Nac
  D_11_r_Nac D_121_r_Nac D_122_r_Nac D_12_r_Nac D_1_r_Nac
  D_41_r_Nac D_41_e_Nac D_42_r_Nac D_4142_re_Nac D_44_r_Nac D_45_r_Nac D_45_e_Nac D_45_re_Nac D_4_r_Nac
D_4_e_Nac D_4_re_Nac B_5n_Nac B_5b_Nac
  D_51_e_Nac D_59_e_Nac D_5_e_Nac
  D_611_e_Nac D_612_r_Nac D_612_e_Nac D_612_er_Nac D_61_r_Nac D_61_e_Nac D_61_er_Nac
  D_623_e_Nac D_62_r_Nac D_62_e_Nac D_62_re_Nac
  D_71_e_Nac D_72_r_Nac D_75_r_Nac D_75_e_Nac D_7_r_Nac D_7_e_Nac B_6b_Nac B_6n_Nac
  AIVPP_P AIVPP_CL_IPPH AIVPP_CL_Seg AIVPP_CL_TCE AIVPP_OIN CCFFEE
  IMALQPRO_B CCFRM ORNC1_B CSIMEMP ORI_B IMALQTER_B CSEFHOG
  CSIMr ORNC2_B ORNC3_B INDNOVIDA1_B ORNC4_B INDNOVIDA2_B
  OTRDOH_B PRIMNOVIDA_B OTPAOH_B
```



```

/ v3 =
O_Npers O_K_1 O_B_2n O_B_2b O_B_3n O_B_3b O_B_2n3n O_B_2b3b
O_D_11_r O_D_121_r O_D_122_r O_D_12_r O_D_1_r
O_D_41_r O_D_41_e O_D_42_r O_D_4142_re O_D_44_r O_D_45_r O_D_45_e O_D_45_re O_D_4_r O_D_4_e O_D_4_re O_B_5n
O_B_5b
O_D_51_e O_D_59_e O_D_5_e
O_D_611_e O_D_612_r O_D_612_e O_D_612_er O_D_61_r O_D_61_e O_D_61_er
O_D_623_e O_D_62_r O_D_62_e O_D_62_re
O_D_71_e O_D_72_r O_D_75_r O_D_75_e O_D_7_r O_D_7_e O_B_6b O_B_6n
O_AIVPP_P O_AIVPP_CI_IPPH O_AIVPP_CI_Seg O_AIVPP_CI_TCE O_AIVPP_OIN O_CCFFEE
O_IMALQPRO_B O_CCFRM O_ORNC1_B O_CSIMEMP O_ORI_B O_IMALQTER_B O_CSEFHOG
O_CSIMr O_ORNC2_B O_ORNC3_B O_INDNOVIDA1_B O_ORNC4_B O_INDNOVIDA2_B
O_OTRDOH_B O_PRIMNOVIDA_B O_OTPAOH_B
.
compute v3 = rnd(v1 - v2 , 0.01).
end repeat.
compute O_AIVPP_P2 = (AIVPP_P_Est - AIVPP_CI_TCE_Est) - AIVPP_P .
compute O_D_4142_r = O_D_4142_re + O_D_41_e.
execute.
delete variables region to O_Npers.

*=====,
* Carga de Ocultaciones de ECV_b2013 vs ECV_b2004.
*=====,
GET FILE='WorkDir\BDOH_b2013_b2004.sav'.
DATASET NAME BDOH WINDOW=FRONT.
DATASET ACTIVATE BDOH.
delete variables w idh13.
rename variables (idh = idh_bdh).

*=====,
* Carga de la BDH e incorporación de las Ocultaciones Agregadas.
*=====,
Get file='BDHDir\BDH.sav' .
DATASET NAME BDH.

DATASET ACTIVATE BDH.
MATCH FILES /FILE=*
/TABLE='BDN'
/BY pais ano.
EXECUTE.
dataset close BDN.

DATASET ACTIVATE BDH.
MATCH FILES /FILE=*
/TABLE='BDOH'
/BY pais ano idh_bdh.
EXECUTE.
dataset close BDOH.

*=====,
* Preparaciones para el cálculo de Indicadores de incongruencia basados en discordancias.
*=====,
SET RNG=MT MTINDEX = -12345.678 .

DATASET ACTIVATE BDH.
FILTER OFF.
USE ALL.
*** SAMPLE .10 .
EXECUTE.
CACHE.

compute filtro = (1=1).
do if missing(Tiphogar).
compute Vseg = 9*10000+ano .
else if (Tiphogar=1 or Tiphogar=2 or Tiphogar=3 or Tiphogar=10).
compute Vseg = 1*10000+ano .
else.
compute Vseg = Tiphogar*10000+ano .
end if.
execute.

*=====,
* 1ª etapa .
*=====,
sort cases by Vseg idh.
*-----,
*----- Indicadores de Incongruencia de la Renta Global -----,
*-----,

```

```

title 'G_001'.
compute filtro = (1=1).
do if (filtro) .
  compute Yvar = B_6n_Bdh / nmh.
  do if ( not missing(GASTOH_epf / nmh) ).
    compute Xvar = GASTOH_epf / nmh.
    compute VVD = Xvar.
  end if.
  compute VVI1 = Yvar.
  compute VVI2 = edad_SP.
  compute VVI3 = G_Consumo_eff.
  compute VVI4 = G_alim_eff.
  compute VVI5 = G_OBDURA_eff.
  compute VVI6 = G_TCAOH_eff.
  compute VVI7 = G_PPEN_eff.
  compute VVI8 = G_SEGVIDA_eff.
  compute VVI9 = PAGOSH_eff.
end if.
execute.
!Completa_Xvar nVars=9 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu TipInd=G
NumInd=001 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu TipInd=G
NumInd=001 Umbral= "-1" .
delete variables filtro Xvar Yvar.

title 'G_002'.
compute filtro = (1=1).
do if (filtro) .
  compute Yvar = B_6n_Bdh / nmh.
  do if ( not missing(GASTOH_epf / nmh) ).
    compute Xvar = GASTOH_eff / nmh.
    compute VVD = Xvar.
  end if.
  compute VVI1 = Yvar.
  compute VVI2 = edad_SP.
  compute VVI3 = G_Consumo_eff.
  compute VVI4 = G_alim_eff.
  compute VVI5 = G_OBDURA_eff.
  compute VVI6 = G_TCAOH_eff.
  compute VVI7 = G_PPEN_eff.
  compute VVI8 = G_SEGVIDA_eff.
  compute VVI9 = PAGOSH_eff.
end if.
execute.
!Completa_Xvar nVars=9 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu TipInd=G
NumInd=002 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu TipInd=G
NumInd=002 Umbral= "-1" .
delete variables filtro Xvar Yvar.

title 'G_003'.
Missing Values riquezaanet_eff (-1, -2).
compute filtro = (1=1).
do if (filtro) .
  compute Yvar = B_6n_Bdh / nmh.
  do if ( not missing(riquezaneff / nmh) ).
    compute Xvar = riquezaanet_eff / nmh.
    compute VVD = Xvar.
  end if.
  compute VVI1 = Yvar.
  compute VVI2 = edad_SP.
  compute VVI3 = actreales_eff.
  compute VVI4 = actfinanc_eff.
  compute VVI5 = vdeuda_eff.
end if.
execute.
!Completa_Xvar nVars=5 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu TipInd=G
NumInd=003 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu TipInd=G
NumInd=003 Umbral= "-1" .
delete variables filtro Xvar Yvar.

title 'G_004'.
Missing Values riquezaabr_eff (-1, -2).
compute filtro = (1=1).
do if (filtro) .
  compute Yvar = B_6n_Bdh / nmh.
  do if ( not missing(riquezabr_eff / nmh) ).

```

```

compute Xvar = riquezaabr_eff / nmh.
compute VVD = Xvar.
end if.
compute VV11 = Yvar.
compute VV12 = edad_SP.
compute VV13 = actreales_eff.
compute VV14 = actfinanc_eff.
compute VV15 = vdeuda_eff.
end if.
execute.
!Completa_Xvar nVars=5 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu Tiplnd=G
NumInd=004 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu Tiplnd=G
NumInd=004 Umbral= "-1" .
delete variables filtro Xvar Yvar.

*-----,
* Corrección de AIVPP_P: Producción relacionada con el Alquiler imputado Vivienda Principal, si el Hogar la ocupa en
propiedad ) -.
*-----,
title 'B2_001'.
Missing Values P2_1_eff P2_5_eff p2_3_eff (-1, -2).
compute filtro = (regteneviv=1).
do if (filtro) .
compute Yvar = AIVPP_P.
do if (P2_1_eff=2).
compute Xvar = P2_5_eff .
compute VVD = Xvar.
end if.
compute VV11 = Yvar.
compute VV12 = edad_SP .
compute VV13 = HH031.
compute VV14 = actreales_eff.
compute VV15 = cuotahip .
end if.
execute.
!Completa_Xvar nVars=5 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu Tiplnd=B2
NumInd=001 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu Tiplnd=B2
NumInd=001 Umbral= "-1" .
delete variables filtro Xvar Yvar.

*-----,
* Corrección de AIVPP_CI_IPPH: AIVPP_CI Intereses de Prestamos Hipotecarios de Propietarios*/.
*-----,
title 'B2_002'.
Missing Values P2_1_eff p2_3_eff p2_12_1_eff p2_13_1_eff (-1, -2).
compute filtro = (regteneviv2=2).
do if (filtro) .
compute Yvar = AIVPP_CI_IPPH.
do if (P2_1_eff=2 and IPPH_B_EFF>=0).
compute Xvar = IPPH_B_EFF .
compute VVD = Xvar.
end if.
compute VV11 = Yvar.
compute VV12 = edad_SP.
compute VV13 = HH031.
compute VV14 = cuotahip .
compute VV15 = IPPH_B.
end if.
execute.
!Completa_Xvar nVars=5 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Sob Tiplnd=B2
NumInd=002 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Sob Tiplnd=B2
NumInd=002 Umbral= "-1" .
delete variables filtro Xvar Yvar.

*-----,
* Corrección de AIVPP_CI_Seg: AIVPP_CI Servicios de Seguros y Planes de Pensiones */.
*-----,
title 'B2_013'.
Missing Values P2_1_eff P5_20b_eff P5_20c_eff p2_12_1_eff p2_13_1_eff P2_5_eff (-1, -2).
compute filtro = (regteneviv=1).
do if (filtro) .
compute Yvar = AIVPP_CI_Seg.
do if (P2_1_eff=2 and IPPH_B_EFF>=0).
compute Xvar = 0.
if (not missing(P5_20b_eff)) Xvar = Xvar + P5_20b_eff .

```

```

        if (not missing(P5_20c_eff)) Xvar = Xvar + P5_20c_eff .
        compute VVD = Xvar.
    end if.
    compute VVI1 = Yvar.
    compute VVI2 = IPPH_B .
    compute VVI3 = G_SEGVIDA_eff.
    compute VVI4 = IPPH_B_EFF.
    compute VVI5 = AIVPP_P.
    compute VVI6 = P2_4_eff .
end if.
execute.
!Completa_Xvar nVars=6 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Sob Tiplnd=B2
NumInd=013 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu Tiplnd=B2
NumInd=013 Umbral= "-1" .
delete variables filtro Xvar Yvar.

title 'B2_003'.
Missing Values P2_1_eff P5_20b_eff P5_20c_eff p2_12_1_eff p2_13_1_eff P2_5_eff (-1, -2).
compute filtro = (regteneviv=1).
do if (filtro) .
    compute Yvar = AIVPP_CI_Seg.
    do if (P2_1_eff=2 and IPPH_B_EFF>=0).
        compute Xvar = IPPH_B_EFF.
        compute VVD = Xvar.
    end if.
    compute VVI1 = Yvar.
    compute VVI2 = IPPH_B .
    compute VVI3 = G_SEGVIDA_eff.
    compute VVI4 = AIVPP_P.
    compute VVI5 = P2_4_eff .
end if.
execute.
!Completa_Xvar nVars=5 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Sob Tiplnd=B2
NumInd=003 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu Tiplnd=B2
NumInd=003 Umbral= "-1" .
delete variables filtro Xvar Yvar.

*-----,
* Corrección de AIVPP_OIN: Otros Impuestos netos sobre la Producción relacionados con el AIVPP -.
*-----,

title 'B2_004'.
Missing Values P2_1_eff P2_5_eff p2_3_eff (-1, -2).
compute filtro = (regteneviv=1).
do if (filtro) .
    compute Yvar = AIVPP_OIN.
    do if (P2_1_eff=2).
        compute Xvar = P2_5_eff .
        compute VVD = Xvar.
    end if.
    compute VVI1 = Yvar.
    compute VVI2 = edad_SP.
    compute VVI3 = HH031.
    compute VVI4 = actreales_eff.
    compute VVI5 = cuotahip .
end if.
execute.
!Completa_Xvar nVars=5 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Sob Tiplnd=B2
NumInd=004 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu Tiplnd=B2
NumInd=004 Umbral= "-1" .
delete variables filtro Xvar Yvar.

*-----,
* Corrección de CCFEE: Consumo de Capital Fijo del Excedente de Explotación -----,
*-----,

title 'B2_005'.
Missing Values P2_1_eff P2_4_eff P2_4_eff p2_3_eff (-1, -2).
compute filtro = (regteneviv=1).
do if (filtro) .
    compute Yvar = CCFEE.
    do if (P2_1_eff=2).
        compute Xvar = 0 .
        if (not (missing(P2_4_eff) or missing(P2_3_eff))) Xvar = P2_4_eff * 1.0329** (ano-P2_3_eff) * Min(1, (ano-P2_3_eff)/50) .
        compute VVD = Xvar.
    end if.
    compute VVI1 = Yvar.

```

```

compute VVI2 = edad_SP.
compute VVI3 = HH031.
compute VVI4 = IPPH_B .
compute VVI5 = AIVPP_P.
compute VVI6 = P2_4_eff .
end if.
execute.
!Completa_Xvar nVars=6 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Sob Tiplnd=B2
NumInd=005 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu Tiplnd=B2
NumInd=005 Umbral= "-1" .
delete variables filtro Xvar Yvar.

*-----,
* Corrección de B_3n: Renta Mixta Neta -----,
*-----,

title 'B3_001'.
compute filtro = (B_3n_Bdh<>0 or (RMIX_B_EFF + IMALQPRO_EFF<>0) and not missing(RMIX_B_EFF) and not
missing(IMALQPRO_EFF)).
do if (filtro) .
compute Yvar = B_3n_Bdh.
do if (RMIX_B_EFF + IMALQPRO_EFF<>0).
compute Xvar = RMIX_B_EFF + IMALQPRO_EFF .
compute VVD = Xvar.
end if.
compute VVI1 = Yvar.
compute VVI2 = edad_SP.
compute VVI3 = RMIX_B .
compute VVI4 = IMALQPRO_B .
end if.
execute.
!Completa_Xvar nVars=4 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu Tiplnd=B3
NumInd=001 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu Tiplnd=B3
NumInd=001 Umbral= "-1" .
delete variables filtro Xvar Yvar.

title 'B3_002'.
Missing Values P3_5_1_eff P3_5_2_eff P3_5_3_eff P3_5_4_eff P3_5_5_eff P3_5_6_eff P3_5_7_eff P3_5_8_eff (-1, -2).
compute XXX=V_CUEYDEP_eff.
do repeat AAA = P3_5_1_eff P3_5_2_eff P3_5_3_eff P3_5_4_eff P3_5_5_eff P3_5_6_eff P3_5_7_eff P3_5_8_eff .
if not missing(AAA) XXX = XXX + AAA .
end repeat.
execute.
compute filtro = ( B_3n_Bdh<>0 or (RMIX_B_EFF + IMALQPRO_EFF<>0) and not missing(RMIX_B_EFF) and not
missing(IMALQPRO_EFF) ).
do if (filtro) .
compute Yvar = B_3n_Bdh.
do if (XXX <>0).
compute Xvar =XXX .
compute VVD = Xvar.
end if.
compute VVI1 = Yvar.
compute VVI2 = edad_SP.
compute VVI3 = RMIX_B .
compute VVI4 = V_CUEYDEP_eff .
compute VVI5 = IMALQPRO_B .
end if.
execute.
!Completa_Xvar nVars=5 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu Tiplnd=B3
NumInd=002 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu Tiplnd=B3
NumInd=002 Umbral= "-1" .
delete variables XXX filtro Xvar Yvar.

*-----,
* Corrección de CCFRM : Consumo de Capital Fijo de la Renta Mixta -----,
*-----,

title 'B3_003'.
Missing Values P4_107_1_eff P4_107_2_eff P4_107_3_eff P4_107_4_eff P4_107_5_eff P4_107_6_eff
P4_108_1_eff P4_108_2_eff P4_108_3_eff P4_108_4_eff P4_108_5_eff P4_108_6_eff (-1, -2).
compute XXX=0.
do repeat AAA = P4_107_1_eff P4_107_2_eff P4_107_3_eff P4_107_4_eff P4_107_5_eff P4_107_6_eff
/ BBB = P4_108_1_eff P4_108_2_eff P4_108_3_eff P4_108_4_eff P4_108_5_eff P4_108_6_eff .
if (not(missing(AAA) or missing(BBB))) XXX = XXX + BBB * 1.0208**(ano-AAA) * Min(1, (ano-AAA)/20) .
end repeat.
execute.

```

```

compute filtro = ( B_3n_Bdh<>0 or (RMIX_B_EFF + IMALQPRO_EFF<>0) and not missing(RMIX_B_EFF) and not
missing(IMALQPRO_EFF) ).
do if (filtro) .
  compute Yvar = CCFRM.
  do if not missing(RMIX_B_EFF) and not missing(IMALQPRO_EFF) .
    compute Xvar =XXX .
    compute VVD = Xvar.
  end if.
  compute VVI1 = Yvar.
  compute VVI2 = edad_SP.
  compute VVI3 = RMIX_B .
  compute VVI4 = V_CUEYDEP_eff .
end if.
execute.
!Completa_Xvar nVars=4 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu Tiplnd=B3
NumInd=003 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu Tiplnd=B3
NumInd=003 Umbral= "-1" .
delete variables XXX filtro Xvar Yvar.

*------.
* Corrección de D_11_r: Sueldos y Salarios -----.
*------.

title 'D1_001'.
compute filtro = ( D_11_r_Bdh > 0 or (RMCA_B_EFF+RNMCA_B_EFF>0) and not missing(RMCA_B_EFF) and not
missing(RNMCA_B_EFF) ).
do if (filtro) .
  compute Yvar = D_11_r_Bdh - ORNC1_B.
  do if (RMCA_B_EFF+RNMCA_B_EFF>0).
    compute Xvar = RMCA_B_EFF+RNMCA_B_EFF .
    compute VVD = Xvar.
  end if.
  compute VVI1 = Yvar.
  compute VVI2 = edad_SP.
  compute VVI3 = RMCA_B .
  compute VVI4 = RNMCA_B .
end if.
execute.
!Completa_Xvar nVars=4 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu Tiplnd=D1
NumInd=001 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu Tiplnd=D1
NumInd=001 Umbral= "-1" .
delete variables filtro Xvar Yvar.

*------.
* Corrección de D.4142.r Intereses (recursos) + Rentas distribuidas de las sociedades (recursos) -----.
*------.

title 'D4_001'.
Missing Values actfinanc_eff (-1, -2).
compute filtro = (D_4142_r_Bdh > 0 or (IMCAP_B_EFF>0 and not missing(IMCAP_B_EFF)) or (actfinanc_eff >0) and not
missing(actfinanc_eff) ).
do if (filtro) .
  compute Yvar = D_4142_r_Bdh.
  do if (IMCAP_B_EFF > 0 or actfinanc_eff >0).
    compute Xvar = IMCAP_B_EFF .
    compute VVD = Xvar.
  end if.
  compute VVI1 = Yvar.
  compute VVI2 = edad_SP.
  compute VVI3 = IMCAP_B.
  compute VVI4 = actfinanc_eff.
end if.
execute.
!Completa_Xvar nVars=4 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu Tiplnd=D4
NumInd=001 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu Tiplnd=D4
NumInd=001 Umbral= "-1" .
delete variables filtro Xvar Yvar.

title 'D4_002'.
Missing Values actfinanc_eff (-1, -2).
compute filtro = (D_4142_r_Bdh > 0 or (IMCAP_B_EFF>0 and not missing(IMCAP_B_EFF)) or (actfinanc_eff >0) and not
missing(actfinanc_eff) ).
do if (filtro) .
  compute Yvar = D_4142_r_Bdh.
  do if (IMCAP_B_EFF > 0 or actfinanc_eff >0).
    compute Xvar = actfinanc_eff .
    compute VVD = Xvar.
  end if.
end if.
execute.
!Completa_Xvar nVars=4 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu Tiplnd=D4
NumInd=001 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu Tiplnd=D4
NumInd=001 Umbral= "-1" .
delete variables filtro Xvar Yvar.

```

```

end if.
compute VVI1 = Yvar.
compute VVI2 = edad_SP_eff .
compute VVI3 = IMCAP_B.
compute VVI4 = IMCAP_B_EFF.
end if.
execute.
!Completa_Xvar nVars=4 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu TipInd=D4
NumInd=002 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu TipInd=D4
NumInd=002 Umbral= "-1" .
delete variables filtro Xvar Yvar.

*-----.
* Corrección de D_41_e: Intereses (empleos) -----.
*-----.

title 'D4_003'.
Missing Values deuotros_eff (-1, -2).
compute filtro = ( D_41_e_Bdh > 0 or deuotros_eff > 0 and not missing(deuotros_eff) ).
do if (filtro) .
compute Yvar = D_41_e_Bdh.
do if (deuotros_eff>0).
compute Xvar = deuotros_eff .
compute VVD = Xvar.
end if.
compute VVI1 = Yvar.
compute VVI2 = dvivpral_eff .
compute VVI3 = deuoprop_eff .
compute VVI4 = vdeuda_eff .
end if.
execute.
!Completa_Xvar nVars=4 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu TipInd=D4
NumInd=003 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu TipInd=D4
NumInd=003 Umbral= "-1" .
delete variables filtro Xvar Yvar.

*-----.
* Corrección de D_44_r: Otras Rentas de Inversión (recursos) -----.
*-----.

title 'D4_004'.
Missing Values allf_eff valor_eff valseg_eff actfinanc_eff (-1, -2).
compute filtro = (D_44_r_Bdh > 0 or actfinanc_eff > 0 and not missing(actfinanc_eff) ).
do if (filtro) .
compute Yvar = D_44_r_Bdh.
do if (allf_eff>0 or valor_eff>0 or valseg_eff>0).
compute Xvar = 0 .
if (not missing(allf_eff)) Xvar = Xvar + allf_eff .
if (not missing(valor_eff)) Xvar = Xvar + valor_eff .
if (not missing(valseg_eff)) Xvar = Xvar + valseg_eff .
compute VVD = Xvar.
end if.
compute VVI1 = Yvar.
compute VVI2 = edad_SP_eff .
compute VVI3 = IMCAP_B.
compute VVI4 = actfinanc_eff.
end if.
execute.
!Completa_Xvar nVars=4 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu TipInd=D4
NumInd=004 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu TipInd=D4
NumInd=004 Umbral= "-1" .
delete variables filtro Xvar Yvar.

*-----.
* Corrección de D_45_re: Rentas de la tierra (rec-emp) -----.
*-----.

title 'D4_006'.
Missing Values IMALQTER_EFF (-1, -2).
compute filtro = (D_45_re_Bdh > 0 or IMALQTER_EFF > 0 and not missing(IMALQTER_EFF) ).
do if (filtro) .
compute Yvar = D_45_re_Bdh.
do if ( IMALQTER_EFF >= 0 ).
compute Xvar = IMALQTER_EFF .
compute VVD = Xvar.
end if.
compute VVI1 = Yvar.
compute VVI2 = edad_SP.

```

```

        compute VVI3 = IMPRO_B_EFF.
        compute VVI4 = IMALQPRO_EFF.
    end if.
    execute.
    !Completa_Xvar nVars=4 segm=Vseg .
    !Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu TipInd=D4
    NumInd=006 Umbral= "-1" .
    !Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu TipInd=D4
    NumInd=006 Umbral= "-1" .
    delete variables filtro Xvar Yvar.

*-----
* Corrección de D_59_e: Otros impuestos corrientes -----
*-----
title 'D5_002'.
Missing Values riquezaabr_eff (-1, -2).
compute filtro = not missing(riquezabr_eff) and (1=1).
do if (filtro) .
    do if ($casenum=1).
        compute Yvar = D_59_e_Bdh.
    else if (ano ne lag(ano)) .
        compute Yvar = 1.
    else.
        compute Yvar = D_59_e_Bdh.
    end if.
    do if not missing(riquezabr_eff) .
        compute Xvar = riquezaabr_eff /* = actfinanc_eff + actreales_eff */ .
        compute VVD = Xvar.
    end if.
    compute VVI1 = Yvar.
    compute VVI2 = actreales_eff.
    compute VVI3 = actfinanc_eff.
end if.
execute.
!Completa_Xvar nVars=3 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu TipInd=D5
NumInd=002 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu TipInd=D5
NumInd=002 Umbral= "-1" .
delete variables filtro Xvar Yvar.

*-----
* Corrección de D_61_r: Cotizaciones sociales netas (Rec) -----
*-----
title 'D6_002'.
Missing Values N_TRABNEG_eff (-1, -2).
compute filtro = (D_61_r_Bdh > 0 or N_TRABNEG_eff > 0 and not missing(N_TRABNEG_eff) ) .
do if (filtro) .
    compute Yvar = D_61_r_Bdh.
    do if (N_TRABNEG_eff >= 0).
        compute Xvar = N_TRABNEG_eff .
        compute VVD = Xvar.
    end if.
    compute VVI1 = Yvar.
    compute VVI2 = edad_SP.
    compute VVI3 = RMCP_B_EFF.
    compute VVI4 = B_3n_Bdh.
    if missing(p4_102_eff) VVI5 = 0.
    if not missing(p4_102_eff) VVI5 = p4_102_eff.
    if missing(p7_12a_eff) VVI6 = 0.
    if not missing(p7_12a_eff) VVI6 = p7_12a_eff .
end if.
execute.
!Completa_Xvar nVars=6 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu TipInd=D6
NumInd=002 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu TipInd=D6
NumInd=002 Umbral= "-1" .
delete variables filtro Xvar Yvar.

*-----
* Corrección de D_62_r: Prest. Soc. distintas de transf. soc. en espec (Rec) -.
*-----
title 'D6_003'.
compute filtro = (D_62_r_Bdh>0 or PRESOC_B_m_eff > 0 and not missing(PRESOC_B_m_eff) ) .
do if (filtro) .
    compute Yvar = D_62_r_Bdh - CSIMEMP + INDNOVIDA1_B.
    do if ( PRESOC_B_m_eff >0 ) .
        compute Xvar = PRESOC_B_m_eff .
        compute VVD = Xvar.
    end if.

```



```

compute VVI1 = Yvar.
compute VVI2 = edad_SP.
compute VVI3 = PSDES_B.
compute VVI4 = PSJYS_B.
compute VVI5 = PSEEL_B.
compute VVI6 = PSEST_B.
compute VVI7 = PSFAM_B.
compute VVI8 = PSVIV_B.
compute VVI9 = PSONC_B.
end if.
execute.
!Completa_Xvar nVars=9 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu TipInd=D6
NumInd=003 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu TipInd=D6
NumInd=003 Umbral= "-1" .
delete variables filtro Xvar Yvar.

*-----
* Corrección de D_62_e: Prest. Soc. distintas de transf. soc. en espec (Emp) -.
*-----
title 'D6_004'.
Missing Values N_TRABNEG_eff (-1, -2).
compute filtro = (D_62_e_Bdh > 0 or N_TRABNEG_eff > 0 and not missing(N_TRABNEG_eff) ) .
do if (filtro) .
  compute Yvar = D_62_e_Bdh.
  do if ( N_TRABNEG_eff >= 0 ).
    compute Xvar = N_TRABNEG_eff .
    compute VVD = Xvar.
  end if.
  compute VVI1 = Yvar.
  compute VVI2 = edad_SP.
  compute VVI3 = RMCP_B_EFF.
  compute VVI4 = B_3n_Bdh.
  if missing(p4_102_eff) VVI5 = 0.
  if not missing(p4_102_eff) VVI5 = p4_102_eff.
  if missing(p7_12a_eff) VVI6 = 0.
  if not missing(p7_12a_eff) VVI6 = p7_12a_eff .
end if.
execute.
!Completa_Xvar nVars=6 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu TipInd=D6
NumInd=004 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu TipInd=D6
NumInd=004 Umbral= "-1" .
delete variables filtro Xvar Yvar.

*-----
* Corrección de D_72_r: Indemnizaciones de seguro no vida (recursos) -----
*-----
title 'D7_001'.
compute filtro = (D_72_r_Bdh > 0 or INDNOVIDA1_B > 0 and not missing(INDNOVIDA1_B) ) .
do if (filtro) .
  compute Yvar = D_72_r_Bdh.
  do if ( INDNOVIDA1_B >= 0 and not missing(INDNOVIDA1_B) ) .
    compute Xvar = INDNOVIDA1_B .
    compute VVD = Xvar.
  end if.
  compute VVI1 = Yvar.
  compute VVI2 = edad_SP.
  compute VVI3 = timpvehic_eff .
  compute VVI4 = valhog_eff .
  compute VVI5 = actreales_eff .
end if.
execute.
!Completa_Xvar nVars=5 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu TipInd=D7
NumInd=001 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu TipInd=D7
NumInd=001 Umbral= "-1" .
delete variables filtro Xvar Yvar.

*-----
* Corrección de D_75_r: Transferencias corrientes diversas (Recursos) -----
*-----
title 'D7_002'.
compute filtro = (D_75_r_Bdh > 0 or (TCDOH_B_EFF > 0 and not missing(TCDOH_B_EFF)) or (OTRDOH_B > 0 and not
missing(OTRDOH_B)) or (P1_11d_eff=1 and anoe=2011) ) .
do if (filtro) .
  compute Yvar = D_75_r_Bdh.
  do if ( not missing(TCDOH_B_EFF) and not missing(OTRDOH_B) ) .

```

```

        compute Xvar = TCDOH_B_EFF+ OTRDOH_B.
        compute VVD = Xvar.
    end if.
    compute VVI1 = Yvar.
    compute VVI2 = edad_SP .
    compute VVI3 = TCDOH_B_EFF .
    compute VVI4 = OTRDOH_B.
end if.
execute.
!Completa_Xvar nVars=4 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu TipInd=D7
NumInd=002 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu TipInd=D7
NumInd=002 Umbral= "-1" .
delete variables filtro Xvar Yvar.

*-----
* Corrección de D_71_e: Primas netas de seguro no vida -----
*-----
title 'D7_003'.
Missing Values P5_23_eff (-1, -2).
compute filtro = (D_71_e_Bdh > 0 or P5_23_eff > 0 and not missing(P5_23_eff) ) .
do if (filtro) .
    compute Yvar = D_71_e_Bdh.
    do if ( P5_23_eff>=0 ).
        compute Xvar = P5_23_eff .
        compute VVD = Xvar.
    end if.
    compute VVI1 = Yvar.
    compute VVI2 = edad_SP.
    compute VVI3 = PRIMNOVIDA_B .
end if.
execute.
!Completa_Xvar nVars=3 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu TipInd=D7
NumInd=003 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu TipInd=D7
NumInd=003 Umbral= "-1" .
delete variables filtro Xvar Yvar.

*-----
* Corrección de D_75_e: Transferencias corrientes diversas (empleos) -----
*-----
title 'D7_004'.
Missing Values P9_4_eff (-1, -2).
compute filtro = ( D_75_e_Bdh>0 or P9_4_eff>0 and not missing(P9_4_eff) ) .
do if (filtro) .
    compute Yvar = D_75_e_Bdh.
    do if ( P9_4_eff>=0 ).
        compute Xvar = P9_4_eff *12 .
        compute VVD = Xvar.
    end if.
    compute VVI1 = Yvar.
    compute VVI2 = edad_SP.
    compute VVI3 = (estcivil_SP_eff=3).
    compute VVI4 = (estcivil_SP_eff=5).
    compute VVI5 = G_Consumo_eff.
    compute VVI6 = G_OBDURA_eff.
    compute VVI7 = G_PPEN_eff.
    compute VVI8 = G_SEGVIDA_eff.
    compute VVI9 = PAGOSH_eff.
end if.
execute.
!Completa_Xvar nVars=9 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu TipInd=D7
NumInd=004 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu TipInd=D7
NumInd=004 Umbral= "-1" .
delete variables filtro Xvar Yvar.

*-----
* Grabación de Datos (SPSS).
*-----
DATASET ACTIVATE BDH.
save OUTFILE='BDHDir\BDHaux1.sav' /COMPRESSED.

*-----
* Correcciones al Final de la 1ª etapa .
*-----
sort cases by pais ano idh.
*-----

```

```

*----- Indicadores de Incongruencia de la Renta Global -----*.
compute filtro = (I=1).
execute.
limputa_minimo_a_no_filtrados indice=G_001 filtro=filtro Umbral="-1".
limputa_minimo_a_no_filtrados indice=G_002 filtro=filtro Umbral="-1".
limputa_minimo_a_no_filtrados indice=G_003 filtro=filtro Umbral="-1".
limputa_minimo_a_no_filtrados indice=G_004 filtro=filtro Umbral="-1".
lsintetiza_indicadores sintesis = G_000 indices = ( G_001 G_002 G_003 G_004 ) Umbral="-1".

*-----*.
* Corrección de AIVPP_P: Producción relacionada con el Alquiler imputado Vivienda Principal, si el Hogar la ocupa en
propiedad ) -.
*-----*. compute filtro = (regteneviv=1).
execute.
limputa_minimo_a_no_filtrados indice=B2_001 filtro=filtro Umbral="-1".
lcalcula_correccion_optima_tasa_lineal partida=AIVPP_P Ox=O_AIVPP_P2
      filtro="(regteneviv=1)"
      x0=B2_001 pesos=w k=90 x1= (G_000).

*-----*.
* Corrección de AIVPP_CI_IPPH: AIVPP_CI Intereses de Prestamos Hipotecarios de Propietarios*/.
*-----*.
compute filtro = (regteneviv2=2).
execute.
limputa_minimo_a_no_filtrados indice=B2_002 filtro=filtro Umbral="-1".
lcalcula_correccion_optima_tasa_lineal partida=AIVPP_CI_IPPH Ox=O_AIVPP_CI_IPPH
      filtro="(regteneviv2=2)"
      x0=B2_002 pesos=w k=90 x1= (G_000).

*-----*. * Corrección de AIVPP_CI_Seg: AIVPP_CI Servicios de Seguros y Planes de Pensiones */.
*-----*.
compute filtro = (regteneviv=1).
execute.
limputa_minimo_a_no_filtrados indice=B2_013 filtro=filtro Umbral="-1".
limputa_minimo_a_no_filtrados indice=B2_003 filtro=filtro Umbral="-1".
lsintetiza_indicadores sintesis = B2_000 indices = ( B2_013 B2_003 ) Umbral="-1".
lcalcula_correccion_optima_tasa_lineal partida=AIVPP_CI_Seg Ox=O_AIVPP_CI_Seg
      filtro="(regteneviv=1)"
      x0=B2_000 pesos=w k=90 x1= (G_000).

*-----*.
* Corrección de AIVPP_OIN: Otros Impuestos netos sobre la Producción relacionados con el AIVPP -.
*-----*.
compute filtro = (regteneviv=1).
execute.
limputa_minimo_a_no_filtrados indice=B2_004 filtro=filtro Umbral="-1".
lcalcula_correccion_optima_tasa_lineal partida=AIVPP_OIN Ox=O_AIVPP_OIN
      filtro="(regteneviv=1)"
      x0=B2_004 pesos=w k=90 x1= (G_000).

*-----*.
* Corrección de CCFEE: Consumo de Capital Fijo del Excedente de Explotación -----*.
*-----*.
compute filtro = (regteneviv=1).
execute.
limputa_minimo_a_no_filtrados indice=B2_005 filtro=filtro Umbral="-1".
lcalcula_correccion_optima_tasa_lineal partida=CCFEE Ox=O_CCFEE
      filtro="(regteneviv=1)"
      x0=B2_005 pesos=w k=90 x1= (G_000).

*-----*.
* Corrección de B_3n: Renta Mixta Neta -----*.
*-----*.
compute filtro = (B_3n_Bdh<>0 or (RMIX_B_EFF + IMALQPRO_EFF<>0) and not missing(RMIX_B_EFF) and not
missing(IMALQPRO_EFF)).
execute.
limputa_minimo_a_no_filtrados indice=B3_001 filtro=filtro Umbral="-1".
limputa_minimo_a_no_filtrados indice=B3_002 filtro=filtro Umbral="-1".
lsintetiza_indicadores sintesis = B3_000 indices = ( B3_001 B3_002 ) Umbral="-1".
lcalcula_correccion_optima_tasa_lineal partida=B_3n_Bdh Ox=O_B_3n
      filtro="(B_3n_Bdh<>0 or (RMIX_B_EFF + IMALQPRO_EFF<>0) and not missing(RMIX_B_EFF) and not
missing(IMALQPRO_EFF))"
      x0=B3_000 pesos=w k=90 x1= (G_000).

*-----*.
* Corrección de CCFRM : Consumo de Capital Fijo de la Renta Mixta -----*.
*-----*.
compute filtro = ( B_3n_Bdh<>0 or (RMIX_B_EFF + IMALQPRO_EFF<>0) and not missing(RMIX_B_EFF) and not
missing(IMALQPRO_EFF) ).
execute.
limputa_minimo_a_no_filtrados indice=B3_003 filtro=filtro Umbral="-1".

```

```

!calcula_correccion_optima_tasa_lineal      partida=CCFRM  Ox=O_CCFRM
      filtro="(B_3n_Bdh<>0 or (RMIX_B_EFF + IMALQPRO_EFF<>0) and not missing(RMIX_B_EFF) and not
missing(IMALQPRO_EFF))"
      x0=B3_003 pesos=w k=90 x1= (G_000).

*-----
* Corrección de D_11_r: Sueldos y Salarios -----
*-----
compute filtro = ( D_11_r_Bdh > 0 or (RMCA_B_EFF+RNMCA_B_EFF>0) and not missing(RMCA_B_EFF) and not
missing(RNMCA_B_EFF) ).
execute.
limputa_minimo_a_no_filtrados indice=D1_001 filtro=filtro Umbral= "-1" .
!calcula_correccion_optima_tasa_lineal      partida=D_11_r_Bdh Ox=O_D_11_r
      filtro="(D_11_r_Bdh > 0 or (RMCA_B_EFF+RNMCA_B_EFF>0) and not missing(RMCA_B_EFF) and not
missing(RNMCA_B_EFF))"
      x0=D1_001 pesos=w k=90 x1= (G_000).

*-----
* Corrección de D.4142_r Intereses (recursos) + Rentas distribuidas de las sociedades (recursos) -----
*-----
Missing Values actfinanc_eff (-1, -2).
compute filtro = (D_4142_r_Bdh > 0 or (IMCAP_B_EFF>0 and not missing(IMCAP_B_EFF)) or (actfinanc_eff >0) and not
missing(actfinanc_eff) ).
execute.
limputa_minimo_a_no_filtrados indice=D4_001 filtro=filtro Umbral= "-1" .
limputa_minimo_a_no_filtrados indice=D4_002 filtro=filtro Umbral= "-1" .
lsintetiza_indicadores sintesis = D4_000 indices = ( D4_001 D4_002 ) Umbral= "-1" .
!calcula_correccion_optima_tasa_lineal      partida=D_4142_r_Bdh Ox=O_D_4142_r
      filtro="(D_4142_r_Bdh > 0 or (IMCAP_B_EFF>0 and not missing(IMCAP_B_EFF)) or (actfinanc_eff >0) and not
missing(actfinanc_eff))"
      x0=D4_000 pesos=w k=90 x1= (G_000).

*-----
* Corrección de D_41_e: Intereses (empleos) -----
*-----
Missing Values deuotros_eff (-1, -2).
compute filtro = ( D_41_e_Bdh > 0 or deuotros_eff > 0 and not missing(deuotros_eff) ).
execute.
limputa_minimo_a_no_filtrados indice=D4_003 filtro=filtro Umbral= "-1" .
!calcula_correccion_optima_tasa_lineal      partida=D_41_e_Bdh Ox=O_D_41_e
      filtro="(D_41_e_Bdh > 0 or deuotros_eff > 0 and not missing(deuotros_eff))"
      x0=D4_003 pesos=w k=90 x1= (G_000).

*-----
* Corrección de D_44_r: Otras Rentas de Inversión (recursos) -----
*-----
Missing Values allf_eff valor_eff valseg_eff actfinanc_eff (-1, -2).
compute filtro = (D_44_r_Bdh > 0 or actfinanc_eff > 0 and not missing(actfinanc_eff) ).
execute.
limputa_minimo_a_no_filtrados indice=D4_004 filtro=filtro Umbral= "-1" .
!calcula_correccion_optima_tasa_lineal      partida=D_44_r_Bdh Ox=O_D_44_r
      filtro="(D_44_r_Bdh > 0 or actfinanc_eff > 0 and not missing(actfinanc_eff) )"
      x0=D4_004 pesos=w k=90 x1= (G_000).

*-----
* Corrección de D_45_re: Rentas de la tierra (rec-emp) -----
*-----
Missing Values IMALQTER_EFF (-1, -2).
compute filtro = (D_45_re_Bdh > 0 or IMALQTER_EFF > 0 and not missing(IMALQTER_EFF) ).
execute.
limputa_minimo_a_no_filtrados indice=D4_006 filtro=filtro Umbral= "-1" .
!calcula_correccion_optima_tasa_lineal      partida=D_45_re_Bdh Ox=O_D_45_re
      filtro="(D_45_re_Bdh > 0 or IMALQTER_EFF > 0 and not missing(IMALQTER_EFF))"
      x0=D4_006 pesos=w k=90 x1= (G_000).

*-----
* Corrección de D_59_e: Otros impuestos corrientes -----
*-----
Missing Values riquezaabr_eff (-1, -2).
compute filtro = not missing(riquezabr_eff) and (1=1).
execute.
limputa_minimo_a_no_filtrados indice=D5_002 filtro=filtro Umbral= "-1" .
!calcula_correccion_optima_tasa_lineal      partida=D_59_e_Bdh Ox=O_D_59_e
      filtro="not missing(riquezabr_eff) and (1=1)"
      x0=D5_002 pesos=w k=90 x1= (G_000).

*-----
* Corrección de D_61_r: Cotizaciones sociales netas (Rec) -----
*-----
Missing Values N_TRABNEG_eff (-1, -2).
compute filtro = (D_61_r_Bdh > 0 or N_TRABNEG_eff > 0 and not missing(N_TRABNEG_eff) ) .

```

```

execute.
limputa_minimo_a_no_filtrados indice=D6_002 filtro=filtro Umbral="-1" .
!calcula_correccion_optima_tasa_lineal partida=D_61_r_Bdh Ox=O_D_61_r
    filtro="(D_61_r_Bdh > 0 or N_TRABNEG_eff > 0 and not missing(N_TRABNEG_eff))"
    x0=D6_002 pesos=w k=90 x1=(G_000).

*-----
* Corrección de D_62_r: Prest. Soc. distintas de transf. soc. en espec (Rec) -.
*-----
compute filtro = (D_62_r_Bdh>0 or PRESOC_B_m_eff > 0 and not missing(PRESOC_B_m_eff)) .
execute.
limputa_minimo_a_no_filtrados indice=D6_003 filtro=filtro Umbral="-1" .
!calcula_correccion_optima_tasa_lineal partida=D_62_r_Bdh Ox=O_D_62_r
    filtro="(D_62_r_Bdh>0 or PRESOC_B_m_eff > 0 and not missing(PRESOC_B_m_eff))"
    x0=D6_003 pesos=w k=90 x1=(G_000).

*-----
* Corrección de D_62_e: Prest. Soc. distintas de transf. soc. en espec (Emp) -.
*-----
Missing Values N_TRABNEG_eff (-1, -2).
compute filtro = (D_62_e_Bdh > 0 or N_TRABNEG_eff > 0 and not missing(N_TRABNEG_eff)) .
execute.
limputa_minimo_a_no_filtrados indice=D6_004 filtro=filtro Umbral="-1" .
!calcula_correccion_optima_tasa_lineal partida=D_62_e_Bdh Ox=O_D_62_e
    filtro="(D_62_e_Bdh > 0 or N_TRABNEG_eff > 0 and not missing(N_TRABNEG_eff))"
    x0=D6_004 pesos=w k=90 x1=(G_000).

*-----
* Corrección de D_72_r: Indemnizaciones de seguro no vida (recursos) -----
*-----
compute filtro = (D_72_r_Bdh > 0 or INDNOVIDA1_B > 0 and not missing(INDNOVIDA1_B)) .
execute.
limputa_minimo_a_no_filtrados indice=D7_001 filtro=filtro Umbral="-1" .
!calcula_correccion_optima_tasa_lineal partida=D_72_r_Bdh Ox=O_D_72_r
    filtro="(D_72_r_Bdh > 0 or INDNOVIDA1_B > 0 and not missing(INDNOVIDA1_B))"
    x0=D7_001 pesos=w k=90 x1=(G_000).

*-----
* Corrección de D_75_r:Transferencias corrientes diversas (Recursos) -----
*-----
compute filtro = (D_75_r_Bdh > 0 or (TCDOH_B_EFF > 0 and not missing(TCDOH_B_EFF)) or (OTRDOH_B > 0 and not
missing(OTRDOH_B)) or (P1_11d_eff=1 and anoe=2011)) .
execute.
limputa_minimo_a_no_filtrados indice=D7_002 filtro=filtro Umbral="-1" .
!calcula_correccion_optima_tasa_lineal partida=D_75_r_Bdh Ox=O_D_75_r
    filtro="(D_75_r_Bdh > 0 or (TCDOH_B_EFF > 0 and not missing(TCDOH_B_EFF)) or (OTRDOH_B > 0 and not
missing(OTRDOH_B)) or (P1_11d_eff=1 and anoe=2011))"
    x0=D7_002 pesos=w k=90 x1=(G_000).

*-----
* Corrección de D_71_e: Primas netas de seguro no vida -----
*-----
Missing Values P5_23_eff (-1, -2).
compute filtro = (D_71_e_Bdh > 0 or P5_23_eff > 0 and not missing(P5_23_eff)) .
execute.
limputa_minimo_a_no_filtrados indice=D7_003 filtro=filtro Umbral="-1" .
!calcula_correccion_optima_tasa_lineal partida=D_71_e_Bdh Ox=O_D_71_e
    filtro="(D_71_e_Bdh > 0 or P5_23_eff > 0 and not missing(P5_23_eff))"
    x0=D7_003 pesos=w k=90 x1=(G_000).

*-----
* Corrección de D_75_e: Transferencias corrientes diversas (empleos) -----
*-----
Missing Values P9_4_eff (-1, -2).
compute filtro = (D_75_e_Bdh>0 or P9_4_eff>0 and not missing(P9_4_eff)) .
execute.
limputa_minimo_a_no_filtrados indice=D7_004 filtro=filtro Umbral="-1" .
!calcula_correccion_optima_tasa_lineal partida=D_75_e_Bdh Ox=O_D_75_e
    filtro="(D_75_e_Bdh>0 or P9_4_eff>0 and not missing(P9_4_eff))"
    x0=D7_004 pesos=w k=90 x1=(G_000).

*-----
* Cálculo de partidas consecuentes tras la 1ª etapa.
*-----
*----- AIVPP_CI: Consumos Intermedios relacionados con el AIVPP: -.
compute AIVPP_CI = AIVPP_CI_IPPH + AIVPP_CI_Seg .
*----- AIVPP_RA: Remuneración de Asalariados relacionados con el AIVPP -.
compute AIVPP_RA = 0 /* por definición es 0, para hogares */.
*----- AIVPP: Alquiler imputado Vivienda Principal, si el Hogar la ocupa en propiedad -.
compute AIVPP = AIVPP_P - AIVPP_CI - AIVPP_RA - AIVPP_OIN .
*>->-N->->- B_2b: Excedente Bruto de Explotación -----

```

```

compute B_2b_Bdh = AIVPP .
*>->-N->->- B_2n: Excedente Neto de Explotación -----,
compute B_2n_Bdh = B_2b_Bdh - CCFEE .
*>->-N->->- B_3b: Renta mixta bruta -----,
compute B_3b_Bdh = B_3n_Bdh + CCFRM .
*>->-N->->- B_2n3n: Excedente de explotación neto / Renta mixta neta -----,
compute B_2n3n_Bdh = B_2n_Bdh + B_3n_Bdh .
*>->-N->->- K_1: Consumo de Capital Fijo -----,
compute K_1_Bdh = CCFEE + CCFRM .
*>->-R->->- B_2b3b: Excedente de explotación neto + Renta mixta neta -----,
compute B_2b3b_Bdh = B_2n3n_Bdh + K_1_Bdh .
execute.

=====,
* 2ª etapa .
=====,
sort cases by Vseg idh.
*-----,
* Corrección de D_121r: Cotizaciones sociales efectivas a cargo de empleadores ---- (en 2ª etapa ) .
*-----,
title 'D1_002'.
compute filtro = ( D_11_r_Bdh > 0 or (RMCA_B_EFF+RNMCA_B_EFF>0) and not missing(RMCA_B_EFF) and not
missing(RNMCA_B_EFF)).
do if (filtro) .
    compute Yvar = D_121_r_Bdh .
    compute Xvar = D_11_r_Bdh .
    compute VVD = Xvar.
    compute VV11 = Yvar.
    compute VV12 = D_122_r_Bdh .
    compute VV13 = D_1_r_Bdh .
end if.
execute.
!Completa_Xvar nVars=3 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu TipInd=D1
NumInd=002 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu TipInd=D1
NumInd=002 Umbral= "-1" .
delete variables filtro Xvar Yvar.

*-----,
* Corrección de D_122r: Cotizaciones sociales imputadas a cargo de empleadores ---- (en 2ª etapa ) .
*-----,
title 'D1_003'.
compute filtro = ( D_11_r_Bdh > 0 or (RMCA_B_EFF+RNMCA_B_EFF>0) and not missing(RMCA_B_EFF) and not
missing(RNMCA_B_EFF) ).
do if (filtro) .
    compute Yvar = D_122_r_Bdh .
    compute Xvar = D_11_r_Bdh / nmh .
    compute VVD = Xvar.
    compute VV11 = Yvar.
    compute VV12 = D_121_r_Bdh .
    compute VV13 = D_1_r_Bdh .
end if.
execute.
!Completa_Xvar nVars=3 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Neg anno=2008 Corr=Ocu TipInd=D1
NumInd=003 Umbral= "-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Neg anno=2011 Corr=Ocu TipInd=D1
NumInd=003 Umbral= "-1" .
delete variables filtro Xvar Yvar.

*-----,
* Correcciones al Final de la 2ª etapa .
=====,
sort cases by pais ano idh.
*-----,
* Corrección de D_121r: Cotizaciones sociales efectivas a cargo de empleadores ---- (en 2ª etapa ) .
*-----,
compute filtro = ( D_11_r_Bdh > 0 or (RMCA_B_EFF+RNMCA_B_EFF>0) and not missing(RMCA_B_EFF) and not
missing(RNMCA_B_EFF)).
execute.
!imputa_minimo_a_no_filtrados indice=D1_002 filtro=filtro Umbral= "-1" .
!calcula_correccion_optima_tasa_lineal partida=D_121_r_Bdh Ox=O_D_121_r
filtro="(D_11_r_Bdh > 0 or (RMCA_B_EFF+RNMCA_B_EFF>0) and not missing(RMCA_B_EFF) and not
missing(RNMCA_B_EFF))"
x0=D1_002 pesox=w k=90 x1= (G_000).

*-----,
* Corrección de D_122r: Cotizaciones sociales imputadas a cargo de empleadores ---- (en 2ª etapa ) .

```

```

*-----,
compute filtro = ( D_11_r_Bdh > 0 or (RMCA_B_EFF+RNMCA_B_EFF>0) and not missing(RMCA_B_EFF) and not
missing(RNMCA_B_EFF) ).
execute.
limputa_minimo_a_no_filtrados indice=D1_003 filtro=filtro Umbral="-1" .
!calcula_correccion_optima_tasa_lineal partida=D_122_r_Bdh Ox=O_D_122_r
filtro="(D_11_r_Bdh > 0 or (RMCA_B_EFF+RNMCA_B_EFF>0) and not missing(RMCA_B_EFF) and not
missing(RNMCA_B_EFF))"
x0=D1_003 pesos=w k=90 x1=(G_000).

*-----,
* Cálculo de partidas consecuentes tras la 2ª etapa
*-----,
*>->-N->->- D_12_r: Cotizaciones sociales a cargo del empleador -----,
compute D_12_r_Bdh = D_121_r_Bdh + D_122_r_Bdh .
*>->-R->->- D_1_r: Remuneración de Asalariados -----,
compute D_1_r_Bdh = D_11_r_Bdh + D_12_r_Bdh .
*>->-N->->- D_4142_re: Saldo (Rec-Emp) de Intereses, Dividendos y Ganacias de Inversiones de Capital -----,
compute D_4142_re_Bdh = D_4142_r_Bdh - D_41_e_Bdh .
**>->-R->->- D_4_re: Rentas de la propiedad (rec-emp) -----,
compute D_4_re_Bdh = D_4142_re_Bdh + D_44_r_Bdh + D_45_re_Bdh .
*>->-N->->- B_5n: Saldo de rentas primarias netas -----,
compute B_5n_Bdh = B_2n3n_Bdh + D_1_r_Bdh + D_4_re_Bdh .
*>->-R->->- B_5b: Saldo de rentas primarias brutas -----,
compute B_5b_Bdh = B_2b3b_Bdh + D_1_r_Bdh + D_4_re_Bdh .

execute.

*=====,
* 3ª etapa .
*=====,
sort cases by Vseg idh.
*-----,
.
* Corrección de D_51_e: Impuestos sobre la renta ----- (en 3ª etapa) .
*-----,

title 'D5_001'.
compute filtro = (1=1) .
do if (filtro) .
compute Yvar = D_51_e_Bdh.
do if ( not missing(B_5n_Bdh) ).
compute Xvar = B_5n_Bdh .
compute VVD = Xvar.
end if.
compute VV11 = Yvar.
compute VV12 = B_2n3n_Bdh.
compute VV13 = D_1_r_Bdh .
compute VV14 = D_4_re_Bdh .
end if.
execute.
!Completa_Xvar nVars=4 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu TipInd=D5
NumInd=001 Umbral="-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu TipInd=D5
NumInd=001 Umbral="-1" .
delete variables filtro Xvar Yvar.

*-----,
* Corrección de D_61_e: Cotizaciones sociales netas (empleos) ----- (en 3ª etapa) .
*-----,

title 'D6_001'.
compute filtro = (D_61_e_Bdh > 0 or D_1_r_Bdh > 0).
do if (filtro) .
compute Yvar = D_61_e_Bdh.
do if not missing(D_11_r_Bdh) .
compute Xvar = D_11_r_Bdh .
end if.
compute VVD = Xvar.
compute VV11 = Yvar.
compute VV12 = D_1_r_Bdh.
compute VV13 = D_12_r_Bdh .
end if.
execute.
!Completa_Xvar nVars=3 segm=Vseg .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2008 Corr=Ocu TipInd=D6
NumInd=001 Umbral="-1" .
!Calculo_de_discordancias Vseg=Vseg idh=idh filtro=filtro Yvar=Yvar Xvar=Xvar Rel=Pos anno=2011 Corr=Ocu TipInd=D6
NumInd=001 Umbral="-1" .

```

delete variables filtro Xvar Yvar.

```

=====
* Correcciones al Final de la 3ª etapa .
=====
sort cases by pais ano idh.
*
* Corrección de D_51_e: Impuestos sobre la renta ----- (en 3ª etapa ).
*
compute filtro = (1=1 ).
execute.
limputa_minimo_a_no_filtrados indice=D5_001 filtro=filtro Umbral= "-1" .
!calcula_correccion_optima_tasa_lineal partida=D_51_e_Bdh Ox=O_D_51_e
      filtro="(1=1)"
      x0=D5_001 pesos=w k=90 x1=(G_000).

*
* Corrección de D_61_e: Cotizaciones sociales netas (empleos) ----- (en 3ª etapa ).
*
compute filtro = (D_61_e_Bdh >0 or D_1_r_Bdh >0).
execute.
limputa_minimo_a_no_filtrados indice=D6_001 filtro=filtro Umbral= "-1" .
!calcula_correccion_optima_tasa_lineal partida=D_61_e_Bdh Ox=O_D_61_e
      filtro="(D_61_e_Bdh >0 or D_1_r_Bdh >0)"
      x0=D6_001 pesos=w k=90 x1=(G_000).

*
* Cálculo de partidas consecuentes tras la 3ª etapa
*
*>->-R->->- D_5_e: Impuestos corrientes sobre la renta, el patrimonio, etc. --.
compute D_5_e_Bdh = D_51_e_Bdh + D_59_e_Bdh .
*>->-R->->- D_61_er: Cotizaciones sociales (empleos-recursos) -----
compute D_61_er_Bdh = D_61_e_Bdh - D_61_r_Bdh .
*>->-R->->- D_62_re: Prest. Soc. distintas de transf. soc. en espec (Rec-Emp) -.
compute D_62_re_Bdh = D_62_r_Bdh - D_62_e_Bdh .
*>->-R->->- D_7_r: Otras transferencias corrientes (Recursos) -----
compute D_7_r_Bdh = D_72_r_Bdh + D_75_r_Bdh .
*>->-R->->- D_7_e: Otras transferencias corrientes (Empleos) -----
compute D_7_e_Bdh = D_71_e_Bdh + D_75_e_Bdh .
*>->-R->->- B_6b: Renta disponible bruta -----
compute B_6b_Bdh = B_5b_Bdh - D_5_e_Bdh - D_61_er_Bdh + D_62_re_Bdh + D_7_r_Bdh - D_7_e_Bdh .
*>->-R->->- B_6n: Renta disponible Neta -----
compute B_6n_Bdh = B_5n_Bdh - D_5_e_Bdh - D_61_er_Bdh + D_62_re_Bdh + D_7_r_Bdh - D_7_e_Bdh .

*----- rpc: Renta Disponible Bruta per cápita -----
compute rpc =B_6b_Bdh / nmh.
*----- wpc: Factor de elevación personal para rentas per capita -----
compute wpc=w*nmh .
*----- rpq1: Renta disponible per capita equivalente Buhmann 0.5 -----
compute rpq1 =B_6b_Bdh / (nmh**0.5).
*----- rpq2: Renta Disponible per capita equivalente OCDE (modificada) -----
compute rpq2 =B_6b_Bdh / (1+0.5*(nmh14 -1)+0.3*(nmh - nmh14)) .
*----- rpq3: Renta Disponible per capita equivalente OCDE / Oxford -----
compute rpq3 =B_6b_Bdh / (1+0.7*(nmh14 -1)+0.5*(nmh - nmh14)).

execute.

=====
* Grabación de Datos (SPSS).
=====
DATASET ACTIVATE BDH.
delete variables UltimaV to Vseg.
compute UltimaV = $sysmis.
execute.
save OUTFILE='BDHDir\BDH.sav' /COMPRESSED.

DATASET ACTIVATE BDH.
AGGREGATE
  /OUTFILE='WorkDir\Valores_optimos_k.sav'
  /BREAK=ano
  /K_AIVPP_P =MEAN(K_AIVPP_P)
  /K_AIVPP_CI_IPPH =MEAN(K_AIVPP_CI_IPPH)
  /K_AIVPP_CI_Seg =MEAN(K_AIVPP_CI_Seg)
  /K_AIVPP_OIN =MEAN(K_AIVPP_OIN)
  /K_CCFFEE =MEAN(K_CCFFEE)
  /K_B_3n_Bdh =MEAN(K_B_3n_Bdh)
  /K_CCFRM =MEAN(K_CCFRM)
  /K_D_11_r_Bdh =MEAN(K_D_11_r_Bdh)
  /K_D_121_r_Bdh =MEAN(K_D_121_r_Bdh)
  /K_D_122_r_Bdh =MEAN(K_D_122_r_Bdh)
  /K_D_4142_r_Bdh =MEAN(K_D_4142_r_Bdh)

```



```

/K_D_41_e_Bdh = MEAN(K_D_41_e_Bdh)
/K_D_44_r_Bdh = MEAN(K_D_44_r_Bdh)
/K_D_45_re_Bdh = MEAN(K_D_45_re_Bdh)
/K_D_51_e_Bdh = MEAN(K_D_51_e_Bdh)
/K_D_59_e_Bdh = MEAN(K_D_59_e_Bdh)
/K_D_61_e_Bdh = MEAN(K_D_61_e_Bdh)
/K_D_61_r_Bdh = MEAN(K_D_61_r_Bdh)
/K_D_62_r_Bdh = MEAN(K_D_62_r_Bdh)
/K_D_62_e_Bdh = MEAN(K_D_62_e_Bdh)
/K_D_72_r_Bdh = MEAN(K_D_72_r_Bdh)
/K_D_75_r_Bdh = MEAN(K_D_75_r_Bdh)
/K_D_71_e_Bdh = MEAN(K_D_71_e_Bdh)
/K_D_75_e_Bdh = MEAN(K_D_75_e_Bdh).

* =====
* Comprobación de resultados de BDH.
* =====
DATASET ACTIVATE BDH .

**>>> comprobaciones de (B.2).
compute dif_B_2n3n_Bdh = B_2n3n_Bdh - (B_2n_Bdh + B_3n_Bdh) .
compute dif_B_2b3b_Bdh = B_2b3b_Bdh - (B_2b_Bdh + B_3b_Bdh) .
compute dif_K_1_Bdh = K_1_Bdh - (B_2b3b_Bdh - B_2n3n_Bdh) .
execute.
DESCRIPTIVES VARIABLES= dif_B_2n3n_Bdh dif_B_2b3b_Bdh dif_K_1_Bdh /STATISTICS=MIN MAX MEAN.

*>>> comprobaciones de (D.1).
compute dif_D_12_r_Bdh = D_12_r_Bdh - (D_121_r_Bdh + D_122_r_Bdh) .
compute dif_D_1_r_Bdh = D_1_r_Bdh - (D_11_r_Bdh + D_12_r_Bdh) .
execute.
DESCRIPTIVES VARIABLES= dif_D_12_r_Bdh dif_D_1_r_Bdh /STATISTICS=MIN MAX MEAN.

*>>> comprobaciones de (D.4).
compute dif_D_4142_re_Bdh = D_4142_re_Bdh - (D_41_r_Bdh - D_41_e_Bdh + D_42_r_Bdh) .
compute dif_D_45_re_Bdh = D_45_re_Bdh - (D_45_r_Bdh - D_45_e_Bdh) .
compute dif_D_4_r_Bdh = D_4_r_Bdh - (D_41_r_Bdh + D_42_r_Bdh + D_44_r_Bdh + D_45_r_Bdh) .
compute dif_D_4_e_Bdh = D_4_e_Bdh - (D_41_e_Bdh + D_45_e_Bdh) .
compute dif_D_4_re_Bdh = D_4_re_Bdh - (D_4142_re_Bdh + D_44_r_Bdh + D_45_re_Bdh) .
execute.
DESCRIPTIVES VARIABLES= dif_D_4142_re_Bdh dif_D_45_re_Bdh dif_D_4_r_Bdh dif_D_4_e_Bdh dif_D_4_re_Bdh /STATISTICS=MIN
MAX MEAN.

*>>> comprobación de la Renta Primaria (B.5).
compute dif_B_5n_Bdh = B_5n_Bdh - (B_2n3n_Bdh + D_1_r_Bdh + D_4_re_Bdh) .
compute dif_B_5b_Bdh = B_5b_Bdh - (B_2b3b_Bdh + D_1_r_Bdh + D_4_re_Bdh) .
execute.
DESCRIPTIVES VARIABLES= dif_B_5n_Bdh dif_B_5b_Bdh /STATISTICS=MIN MAX MEAN.

*>>> comprobación de la Renta Primaria (D.5).
compute dif_D_5_e_Bdh = D_5_e_Bdh - (D_51_e_Bdh + D_59_e_Bdh) .
execute.
DESCRIPTIVES VARIABLES= dif_D_5_e_Bdh /STATISTICS=MIN MAX MEAN.

*>>> comprobación de las Cotizaciones Sociales (D.61.er).
compute dif_D_61_r_Bdh = D_61_r_Bdh - (D_612_r_Bdh) .
compute dif_D_61_e_Bdh = D_61_e_Bdh - (D_611_e_Bdh + D_612_e_Bdh) .
compute dif_D_612_er_Bdh = D_612_er_Bdh - (D_612_e_Bdh - D_612_r_Bdh) .
compute dif_D_61_er_Bdh = D_61_er_Bdh - (D_61_e_Bdh - D_61_r_Bdh) .
compute dif_D_61_er_Bdh = D_61_er_Bdh - (D_611_e_Bdh + D_612_er_Bdh) .
compute dif_D_62_re_Bdh = D_62_re_Bdh - (D_62_r_Bdh - D_62_e_Bdh) .
execute.
DESCRIPTIVES VARIABLES= dif_D_61_r_Bdh dif_D_61_e_Bdh dif_D_612_er_Bdh dif_D_61_er_Bdh dif_D_61_er_Bdh
dif_D_62_re_Bdh /STATISTICS=MIN MAX MEAN.

*>>> comprobación de la (D.7).
compute dif_D_7_r_Bdh = D_7_r_Bdh - (D_72_r_Bdh + D_75_r_Bdh) .
compute dif_D_7_e_Bdh = D_7_e_Bdh - (D_71_e_Bdh + D_75_e_Bdh) .
execute.
DESCRIPTIVES VARIABLES= dif_D_7_r_Bdh /STATISTICS=MIN MAX MEAN.

*>>> comprobación de la Renta Disponible del Hogar (B.6b).
compute dif_B_6n_Bdh = B_6n_Bdh - (B_5n_Bdh - D_5_e_Bdh - D_61_er_Bdh + D_62_re_Bdh + D_7_r_Bdh - D_7_e_Bdh) .
compute dif_B_6b_Bdh = B_6b_Bdh - (B_5b_Bdh - D_5_e_Bdh - D_61_er_Bdh + D_62_re_Bdh + D_7_r_Bdh - D_7_e_Bdh) .
execute.
DESCRIPTIVES VARIABLES= dif_B_6n_Bdh dif_B_6b_Bdh /STATISTICS=MIN MAX MEAN.

*>>> comprobación de los missings de los indicadores de incongruencia.
FRECUENCIAS VARIABLES=B2_001 B2_002 B2_013 B2_003 B2_004 B2_005 B3_001 B3_002 B3_003 D1_001
D1_002 D1_003 D4_001 D4_002 D4_003 D4_004 D4_006 D5_001 D5_002 D6_001 D6_002 D6_003 D6_004
D7_001 D7_002 D7_003 D7_004 G_001 G_002 G_003 G_004 G_000 B2_000 B3_000 D4_000
/FORMAT=NOTABLE

```

```
/PERCENTILES=1.0 5.0 10.0 25.0 50.0 75.0 90.0 95.0 99.0 99.5 99.9  
/STATISTICS=MINIMUM MAXIMUM  
/ORDER=ANALYSIS.
```

```
*=====,  
* Salida.  
*=====,  
NEW FILE.  
DATASET CLOSE BDH.  
OUTPUT SAVE NAME=salida  
OUTFILE='LogDir\040e_Corrige_ocultaciones.spv'  
LOCK=NO.  
OUTPUT CLOSE salida.
```